

UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN



**Sistema de visión computacional basada en reconocimiento facial para
identificación de personas dirigido a los miembros de la organización
regional de ciegos del Perú – Chiclayo**

**TESIS PARA OPTAR EL TÍTULO DE
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

AUTOR

Mirella Jazmin Zuloeta Lopez

ASESOR

Huiler Juanito Mera Montenegro

<https://orcid.org/0000-0001-6830-5415>

Chiclayo, 2025

**Sistema de visión computacional basada en reconocimiento facial
para identificación de personas dirigido a los miembros de la
organización regional de ciegos del Perú – Chiclayo**

PRESENTADA POR
Mirella Jazmin Zuloeta Lopez

A la Facultad de Ingeniería de la
Universidad Católica Santo Toribio de Mogrovejo
para optar el título de

INGENIERO DE SISTEMAS Y COMPUTACIÓN

APROBADA POR

Ricardo David Iman Espinoza
PRESIDENTE

Miguel Angel Diaz Espino
SECRETARIO

Huiler Juanito Mera Montenegro
VOCAL

Dedicatoria

Dedico principalmente a Dios por permitirme llegar hasta este momento tan determinante de mi formación profesional.

A mi madre quien durante este tiempo me brindó su amor incondicional en todo momento confiando en mis capacidades de lograr este sueño. Gracias por todo madre.

A mis hermanos y sobrinos por su apoyo moral durante este viaje académico.

Agradecimientos

A Dios por brindarme salud y no dejar que me rinda ante situaciones difíciles.

Al Ing. Oscar Alex Serquén Yparraguirre por el asesoramiento y tiempo invertido para que pueda concluir exitosamente el trabajo.

A la asociación ORCIP y a todas las personas que forman parte de ella por proporcionar el tiempo y los recursos necesarios para ejecutar esta investigación.

Finalmente, agradezco a mis amigos quienes sin esperar nada a cambio me transmitieron sus conocimientos y estuvieron en momentos de alegría y tristeza.

ARTICULO

INFORME DE ORIGINALIDAD

21 %
INDICE DE SIMILITUD

20 %
FUENTES DE INTERNET

6 %
PUBLICACIONES

11 %
TRABAJOS DEL
ESTUDIANTE

FUENTES PRIMARIAS

1	hdl.handle.net Fuente de Internet	4 %
2	tesis.usat.edu.pe Fuente de Internet	2 %
3	repositorioinstitucional.buap.mx Fuente de Internet	1 %
4	Submitted to Corporación Universitaria Minuto de Dios, UNIMINUTO Trabajo del estudiante	1 %
5	renati.sunedu.gob.pe Fuente de Internet	1 %
6	Submitted to Consorcio CIXUG Trabajo del estudiante	1 %
7	Submitted to Shepherd University Trabajo del estudiante	<1 %
8	digital.library.txstate.edu Fuente de Internet	<1 %

Índice

Resumen	6
Abstract	7
Introducción	8
Revisión de literatura	10
Materiales y métodos	16
Resultados y discusión	19
Conclusiones	32
Recomendaciones.....	32
Referencias.....	33
Anexos.....	38

Resumen

En el presente estudio se desarrolló un sistema de visión computacional basada en reconocimiento facial para identificación de personas con el propósito de que las personas con discapacidad visual se sientan autónomas al utilizar la herramienta además de no ser dependientes de su entorno. Para lograr el objetivo principal, se realizó una comparativa de arquitecturas existentes basados en criterios de rendimiento, portabilidad y costos; siendo Facenet la arquitectura idónea para la construcción del algoritmo. Asimismo, se evaluó los modelos existentes de reconocimiento facial mediante procesamiento de imágenes, siendo Facenet el modelo óptimo con 98% de accuracy. En cuanto a la validación del modelo, a través de las métricas de visión computacional en precisión, exactitud y F1; los resultados fueron 100%, 97% y 96% respectivamente. Para la validación del funcionamiento, se encuestaron a 13 personas de la asociación donde el 81% consideraron completamente en cuanto a facilidad de uso percibida y el 95% calificaron la misma categoría en cuanto a utilidad percibida a través del modelo de aceptación tecnológica TAM. Cabe resaltar que la selección de la arquitectura, construcción y validación del modelo se elaboraron dentro del proceso de preparación de datos e ingeniería de atributos proporcionado por Google. Asimismo, para la integración de todo el sistema de visión computacional se desarrolló utilizando Scrum con duración de 3 sprints. El despliegue del modelo se construyó utilizando el lenguaje Python, exportado con TensorFlow Lite, y para la aplicación móvil, se implementó bajo el lenguaje Dart con Flutter Framework para aplicaciones móviles híbridas.

Palabras clave: Visión computacional, reconocimiento facial, discapacidad visual, aplicación móvil, Facenet.

Abstract

In the present study, a computer vision system based on facial recognition was developed for identifying people with the purpose of making people with visual disabilities feel autonomous when using the tool in addition to not being dependent on their environment. To achieve the main objective, a comparison of existing architectures was carried out based on performance, portability and cost criteria; Facenet being the ideal architecture for the construction of the algorithm. Likewise, existing facial recognition models were evaluated through image processing, with Facenet being the optimal model with 98% accuracy. Regarding the validation of the model, through computer vision metrics in precision, accuracy and F1; The results were 100%, 97% and 96% respectively. To validate the operation, 13 people from the association were surveyed, where 81% considered it completely in terms of perceived ease of use and 95% rated the same category in terms of perceived usefulness through the TAM technological acceptance model. It should be noted that the selection of the architecture, construction and validation of the model were developed within the data preparation and attribute engineering process provided by Google. Likewise, for the integration of the entire computer vision system, it was developed using Scrum with a duration of 3 sprints. The model deployment was built using the Python language, exported with TensorFlow Lite, and for the mobile application, it was implemented under the Dart language with the Flutter Framework for hybrid mobile applications.

Keywords: Computer vision, facial recognition, visual impairment, mobile application, Facenet.

Introducción

Según la Organización Mundial de la Salud (OMS), aproximadamente 2200 millones de personas en todo el mundo presentan discapacidad visual o ceguera, donde al menos 1000 millones padecen de discapacidad visual que podría haberse prevenido o aplicado un tratamiento [1]. En la actualidad, el movimiento independiente de las personas que presentan esta deficiencia se encuentra entre los problemas más importantes de la sociedad que deben abordarse [2]. Lamentablemente, a menudo son excluidos de la vida social, económica y política en su entorno por discriminación directa o falta de inclusión. Además, estas personas están expuestas a diversas actividades de la vida cotidiana que involucran el desplazamiento de un lugar a otro y tener que reconocer a las personas que les rodean [3].

De acuerdo con la Fundación ONCE para la Solidaridad con Personas Ciegas de América Latina (FOAL), aproximadamente el 90% de personas con dicha discapacidad se concentran en países de ingresos bajos, donde solo el 10% reciben un apoyo para orientación y movilidad [5]. En Perú, según el Instituto Nacional de Estadística e Informática (INEI), el 10.4% del total de la población, que representa aproximadamente 3 millones, tienen discapacidad con un 57% de mujeres y el 43% restante de hombres, lo que significa que uno de cada 10 peruanos padece de ello [6]. En Chiclayo se encuentra la Organización Regional de Ciegos del Perú (ORCIP) que forma parte del FOAL, el cual el 90% de los miembros de la asociación presentan ceguera total, lo cual son dependientes de otras personas para que puedan realizar sus actividades diarias a diferencia del 10% que presentan ceguera parcial donde a pesar de ello, también se les dificulta ver. No obstante, de las personas que presentan ceguera total, el 20% fue causado por accidentes automovilísticos; el 10% presentan ceguera total de nacimiento, es decir, no se imaginan cómo es la vida, la forma o el color; y el resto es por causa de enfermedades como glaucoma que no fue detectada a tiempo.

La principal causa de discapacidad visual puede ser congénita o como resultado de accidentes, envejecimiento o enfermedades [9]. Además, las causas varían según el país, la disponibilidad, la accesibilidad y el conocimiento público de los servicios de atención oftalmológica. Por ejemplo, los países con ingresos bajos y medianos tienen mayor probabilidad en deterioro de la visión debido a cataratas. Por otro lado, enfermedades como retinopatía diabética, glaucoma y degeneración macular relacionada con la edad son más comunes en los países de ingresos altos.

En consecuencia, debido a las deficiencias que presentan este sector de personas con discapacidad, se sienten limitados en realizar actividades cotidianas; por ejemplo, realizar independientemente las compras o no poder reconocer quién está a su alrededor. Con respecto a lo último, de acuerdo a las encuestas realizadas a las personas de la asociación ORCIP, el 90% precisan que se sienten vulnerables día a día dentro de su entorno al no poder reconocer cuando sienten la presencia de alguien, lo cual ese temor les genera inseguridades y aislamiento social además de que no cuentan con algún medio o mecanismo que les garantice una identificación de las personas. Asimismo, los encuestados enfatizan de que ponen en riesgo su vida en ocasiones donde sus familiares no se encuentran dentro de sus hogares y sienten el miedo ante algún atentado en sus viviendas por no poder ver y que alguien extraño se hagan pasar como sus convivientes. Con todo lo mencionado, las personas con discapacidad visual se sienten perdidas, excluidas, inseguras y ansiosas afectando así su calidad de vida [2].

Ante esta realidad, se planteó la siguiente pregunta ¿De qué manera se puede apoyar a los miembros de la Organización regional de ciegos del Perú – Chiclayo para que puedan identificar personas en un entorno? Es así como, el trabajo se justifica en el ámbito científico porque servirá como base para otros estudios que involucren a personas con discapacidad visual, ya que este grupo de personas es relegado en investigaciones especialmente tecnológicas, además, la importancia de sistemas de visión computacional hace de este campo un área interesante para estudiar e investigar. En lo tecnológico, el uso de las técnicas de inteligencia artificial por medio de dispositivos móviles garantiza el reconocimiento de personas en tiempo real, lo que permite que las personas que presentan deficiencia visual identifiquen rápidamente con quién están interactuando en ese momento sin ninguna preocupación. Es por ello que en este trabajo de investigación planteó un producto, específicamente un sistema de visión computacional utilizando redes neuronales para el reconocimiento facial, además de tecnologías que permitan el desarrollo integrado del sistema.

Para dar respuesta a la formulación del problema, se definió como objetivo general implementar un sistema de visión computacional basada en reconocimiento facial para identificación de personas dirigido a los miembros de la Organización regional de ciegos del Perú – Chiclayo. Asimismo, se plantearon cuatro objetivos específicos, de las cuales son determinar la arquitectura del modelo que permita dar soporte al sistema de visión computacional para la ejecución del algoritmo. Luego, como segundo objetivo, es construir el modelo de reconocimiento facial que se usará para el desarrollo del sistema

de visión computacional. Por consiguiente, se propuso obtener una validación aprobatoria del modelo de reconocimiento facial a través de métricas para visión computacional, y como ultimo objetivo es validar el funcionamiento del sistema de visión computacional con respecto a los criterios de usabilidad y utilidad percibida por los miembros de la Organización regional de ciegos del Perú – Chiclayo.

Revisión de literatura

Para el desarrollo de la presente investigación se hizo una revisión exhaustiva de diversas fuentes, las cuales sustentan y afianzan la solución propuesta.

Antecedentes

Esta investigación ha tomado en cuenta los siguientes antecedentes:

Antecedentes internacionales

Yung [11], realizó una investigación donde su problemática consistía que las personas que presentan discapacidad visual eran dependientes, lo que les resultaba difícil poder reconocer objetos a su alrededor, motivo por lo que desarrolló un aplicativo móvil de detección de objetos a través de las tecnologías Java y Android Studio. El valor agregado del producto es el acceso al usuario con dicha discapacidad visual hacer uso de su propio teléfono inteligente de manera autónoma, dándole así más confianza. Por ello, el autor concluyó que desarrollar este tipo de tecnologías a través de una aplicación móvil le impartió una amplia gama de conocimientos tratando de que sea útil para el público objetivo.

Goldberg [12], consideró como problemática la dificultad de las personas para desplazarse de manera segura, lo cual le permitió desarrollar un marco ciberfísico móvil para que los usuarios que presentan discapacidad visual puedan interactuar, desde el entorno virtual con el mundo real. El valor agregado del trabajo fue el uso de IoT junto con la inteligencia artificial llegando a concluir que se logró una orientación y navegación adecuada para las personas con discapacidad visual.

Márquez et al. [13], se basaron en la dificultad que presentan las personas con dicha discapacidad al no poder reconocer personas, por lo que el sistema que decidieron implementar incluyó la integración de la inteligencia artificial, donde su función principal es poder reconocer los rostros de las personas con las que el usuario interactúa a diario.

El valor adicional de la investigación es que la propuesta utiliza un modelo híbrido de memorias asociativas para el entrenamiento de las personas a partir de diferentes expresiones faciales llegando el autor a concluir que este tipo de reconocimientos ataca directamente el desempeño de los algoritmos al momento de reconocer a las personas en condiciones reales, motivo por lo que hizo uso del algoritmo idóneo.

Antecedentes nacionales

Carreño [14], se enfocó en la accesibilidad que deberían tener las personas invidentes para que puedan utilizar las tecnologías, motivo por lo que implementó una aplicación móvil en Android siendo su valor agregado el acceso al contenido de noticias dentro de la herramienta tecnológica para el público objetivo. Con lo mencionado, el autor concluye que basándose en principios se mejoró accesibilidad para las personas con discapacidad visual. Se escogió dicho antecedente debido a que el autor hace uso de herramienta móvil y sigue ciertos patrones de accesibilidad que se toman de referencia para la presente investigación.

Herencia y Heredia [15], se enfocaron en los inconvenientes que presentan las personas con deficiencia visual en el acceso a su información universitaria por medios digitales. En base a ello, los autores consideraron implementar una interfaz intermediaria con la intranet de la propia universidad para el uso de la personas con problemas visuales. El valor agregado de la solución fueron las herramientas incorporadas de asistencia artificial, lectores de pantalla, servicios de Microsoft considerando el desarrollo de manera híbrida para dispositivos Android e iOS, y la ayuda del campus universitario. Gracias a ello, el autor concluye que resultó más fácil para las personas con dicha discapacidad el uso del aplicativo por medio del dispositivo móvil.

Ríos y Guerrero [16], se basaron en la problemática donde las personas que presentan dificultad motora no pueden generar ordenes de desplazamiento, razón por la cual, los autores decidieron optar por las técnicas de inteligencia artificial tales como procesamiento de imágenes y robótica. Por ello, consideraron que el valor agregado fue el diseño e implementaron un sistema inteligente capaz de generar órdenes de desplazamiento por medio de las vistas de las personas que presentan deficiencia visual, llegando a la conclusión de que el software inteligente influyó de manera adecuada para el reconocimiento de órdenes de desplazamiento por medio de movimientos de las personas que presenten dicha discapacidad.

Antecedentes locales

Sánchez [17], centró su problemática en los obstáculos que tienen los estudiantes con deficiencia visual en el proceso de enseñanza y aprendizaje con el sistema braille. Es así que desarrolló un aplicativo móvil de reconocimiento de palabras siendo su valor agregado el uso de un dispositivo táctil para el aprendizaje braille. Con lo implementado, el autor concluye que se había acortado el proceso de enseñanza braille de la manera tradicional además de que el sistema web se podía usar en conjunto tanto el docente como estudiante para el monitoreo del aprendizaje.

Fernández [18], centró su problemática en la dificultad de desplazamiento que tienen las personas con dicha discapacidad a diario, sintiéndose como una carga para su familia debido a la limitación que ellos presentan, además que son discriminados en la sociedad. En base a ello, el autor consideró implementar un bastón geolocalizador para las personas invidentes siendo su valor agregado el uso de tecnologías integradas como sensores, geolocalizadores, Arduino, motores vibradores y Android para el desarrollo móvil. Con ello, el autor concluye que el uso del bastón inteligente facilita el desplazamiento de las personas invidentes a través del geolocalizador.

Sandoval [19], consideró como problemática en las necesidades de desarrollo educativo para las personas que tienen discapacidad visual ya sea de manera parcial o total. Con lo resaltado, es que empleó construir una herramienta que apoye en el proceso de aprendizaje, siendo su valor adicional el uso de locución automática basado en lectoescritura específicamente para los niños invidentes. Es así que, con su propuesta, concluye que tuvo grandes resultados logrando que los niños presenten calificaciones aprobatorias elevando su promedio de notas.

Bases teórico científicas

Discapacidad visual

Forma parte de la condición humana y es definido como una disminución de la capacidad de la visión o de las funciones del sistema visual. Se clasifica dentro de la baja visión y ceguera [20]. Dentro de las condiciones posteriores que pueden llegar a presentar las personas con dicha deficiencia tiene que ver relacionada con depresión, ansiedad u otros problemas psicológicos donde prevalecen las personas de edad adulta [21]. Para el trabajo de tesis se tomó como objeto de estudio a las personas que presenten dicha

discapacidad dentro de la asociación ORCIP debido a que en su mayoría se encuentran limitados a realizar actividades sociales y personales de manera independiente y con la propuesta les ayudó a tener una mayor autonomía a utilizar dicha herramienta.

Causas de la discapacidad visual

Las causas del deterioro de la visión están relacionadas con glaucoma, degeneración macular, diabetes, traumatismos, retinopatía del prematuro, errores de refracción no corregidos, cataratas, daños en la corteza visual, prevalencia de la neuropatía óptica glaucomatosa y opacidad de las corneas, u otros factores hereditarios o congénitos [20]. Además, se establece de que las causas son originadas de manera congénita grave, envejecimiento de la población o accidentes en el transcurso de su vida [3]. En el trabajo de investigación, se consideró que el uso de toda la aplicación será para el público que presenten ceguera total por diversas causas, incluidas las personas que son invidentes desde nacimiento.

Tratamiento de la discapacidad visual

Existen técnicas como el autocuidado donde son efectivas para mejorar la salud mental, así como su calidad de vida física y mayor autoeficacia. Asimismo, la rehabilitación de la visión es ideal para mejorar síntomas psiquiátricos donde involucra servicios clínicos. Finalmente, otras opciones consideradas para las personas con dicha discapacidad son las intervenciones conductuales cognitivas e intervenciones de atención escalonadas que van relacionadas a las personas que presentan síntomas de depresión y ansiedad [21]. Para el caso de la asociación donde se ha trabajado la tesis, se pudo obtener información de que ellos optaron por el ultimo tipo de tratamiento mencionado, que son las intervenciones conductuales y cognitivas ya que, a día de hoy, han sabido lidiar con la depresión y asimilar su discapacidad.

Inteligencia artificial

Es un término orientado al comportamiento inteligente con capacidad de percibir, razonar, aprender, comunicarse y operar en una variedad de entornos. Su propósito es poder desarrollar máquinas que realicen actividades tal cual como los humanos [22]. Para

el presente trabajo se empleó por este campo de investigación ya que presenta una gama de clasificaciones además de soluciones inteligentes como lo es una gran herramienta de asistencia para personas con discapacidad visual que requieren de este avance tecnológico para una mayor autonomía reflejado así en un producto de aplicación móvil.

Visión computacional

Forma parte de la inteligencia artificial basada no solo para la creación de programas, huellas dactilares y cualquier objeto en particular, sino también para el procesamiento y clasificación de imágenes de manera autónoma [23]. Uno de los usos que forma parte de la visión computacional es el uso de Reconocimiento de Caracteres Ópticos (OCR), el cual es una técnica que genera texto a través de una imagen [24]. No obstante, las computadoras tienen ciertas limitaciones, ya que los programas a menudo deben reconocer ciertos patrones a partir de objetos, imágenes predeterminados, lo cual es un desafío [25]. En este trabajo de investigación es optado por esta área ya que se encuentra dentro de las capacidades que permita reconocer e identificar rostros, así como el procesamiento y análisis de imágenes.

Reconocimiento facial

Es una aplicación que forma parte del reconocimiento de patrones siendo aplicado en la seguridad, identificación de personas, entre otros [26]. Un sistema de reconocimiento de rostros presenta tres módulos: Preprocesamiento, criterios de selección y módulo de clasificación. Los pasos de preprocesamiento incluyen la segmentación, proceso para encontrar datos faciales en una imagen donde incluye la extracción de características de los datos de la cara segmentada como ojos, nariz, labios, etc. [27]. Es así que en el trabajo de tesis se optó por esta técnica claramente clasificada como solución el cual pasa por procesos que son la detección del rostro, extracción de las características del rostros, luego la búsqueda de esas características entre todos los rostros almacenados y finalmente pasa por el reconocimiento facial mediante una comparación y verificación.

Aplicación móvil

Es un software capaz de mejorar la productividad de los usuarios a través de los smartphones. Una aplicación móvil está incluida en todos los dispositivos desde hace tiempo. En la actualidad, existen aplicaciones de todo tipo, color y forma; pero tiempo atrás, solamente era usado para fines básicos como el uso de alarmas, calculadoras, calendarios o correo. A diferencia de una aplicación web, este tipo de aplicaciones se deben descargar e instalar antes de su debido uso, donde no sucede en una aplicación web porque simplemente teniendo internet y un navegador se puede acceder rápidamente a ella [28]. Para el presente trabajo de investigación se consideró idóneo el ser elaborado en un aplicativo móvil debido a su accesibilidad directa con el usuario para que use la aplicación y ya no este accediendo a un navegador constantemente ya que lo tendría instalado en su propio dispositivo.

Metodología de desarrollo de software

Es la determinación del método ideal para la implementación de software de calidad que puede satisfacer a los usuarios. Asimismo, es considerado parte esencial dentro de la ingeniería de software. Existen diferentes clasificaciones de metodologías tales como metodologías clásicas, metodologías orientadas a objetos, metodologías ágiles, metodologías formales y metodologías para la web, entre otros [29]. Para este informe se consideró usar un marco ágil iterativo SCRUM como opción idónea para que el usuario participe de manera recurrente y nos brinde retroalimentaciones durante la implementación.

SCRUM

Es un marco de trabajo para la implementación de software complejos que puedan ser entregados en un tiempo debido de forma sencilla. No es metodología, sino un marco incluido dentro de la metodología de desarrollo de software ágil, donde lo hace dispuesto a que se cree un software eficiente; en otras palabras, es el subconjunto de dicha metodología. SCRUM tiene su foco principal en las actividades de toda la gestión de un proyecto lo cual tiene como objetivo poder satisfacer a sus clientes entregando rápidamente el producto en un corto plazo [30].

Preparación de datos e ingeniería de atributos en Aprendizaje Automático (AA)

Es una secuencia de pasos del aprendizaje automático proporcionado por Google que nos ayuda a encontrar patrones en los datos que usamos para hacer predicciones sobre otros nuevos. Para que estas predicciones sean correctas, necesitamos crear un conjunto de datos y transformarlos adecuadamente. Gracias a la ingeniería de atributos, nos ayuda en los modelos a comprender los conjuntos de datos de la forma en que lo realiza un humano [33].

Materiales y métodos

Tipo de investigación

La investigación aplicada consiste en plantear una idea propuesta en base a nuestro conocimiento para que posteriormente se solucione en un problema específico [34]. Por ello, considerando que el informe presenta justificación, solución tecnológica, sumado a ello la comparación, evaluación, interpretación, determinar precedentes y la causalidad y sus implicaciones; la investigación es de tipo aplicada.

Métodos de investigación

Los métodos en el informe son seis y se describen en la Tabla I:

Método	Sustento por el cual será empleado en la investigación
Analítico	Se realizó un estudio y análisis de la problemática actual que presentan las personas que presenta discapacidad visual.
Analítico sintético	Se realizó el desarrollo de la introducción donde cada párrafo posee su propia estructura, pero en conjunto componen el planteamiento de la situación problemática de las personas con discapacidad visual.
Deductivo	Se realizó la introducción partiendo de lo general a lo particular como en la búsqueda de estadísticas internacionales en la problemática de discapacidad visual en diferentes países.
Análisis de la literatura	Se realizaron búsquedas de antecedentes donde se hayan implementado soluciones similares para personas con discapacidad visual y así dar argumento al trabajo de investigación.
Estudio de caso	Solo se trabajó en el entorno de la asociación ORCIP Chiclayo.
Implementación	Se implementó una solución tecnológica orientada para las personas con discapacidad visual.

Técnicas e instrumentos de recolección de datos

Se consideró a las entrevistas y encuestas, las cuales son consignadas en la Tabla II:

TABLA II
TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

Técnicas	Instrumentos	Elementos de la población	Propósito
Entrevista	Guía de entrevista y grabador de audio (<i>ver anexo N°02 y anexo N°03</i>)	Presidente de la asociación y miembros de la Organización Regional de Ciegos del Perú	Se empleó para recoger información acerca del problema de la discapacidad visual que tienen los miembros de la asociación.
Encuesta	Cuestionario (<i>ver anexo N°04</i>)	Miembros de la Organización Regional de Ciegos del Perú	Se empleó para verificar la usabilidad y utilidad del sistema de visión computacional a implementar a través de datos válidos y confiables.

Procedimientos

Metodología de desarrollo

En el informe se inició con la secuencia de pasos del Proceso de preparación de datos y de ingeniería de atributos para la inteligencia artificial, seleccionando y aplicando el algoritmo idóneo [32]. Además, se llevó a cabo la elaboración de tres sprints considerando las fases del marco de referencia SCRUM; esto fue con el propósito de gestionar el proyecto e iterarlo rápidamente para que sea adaptable a futuras necesidades [31], [35]. Con todo ello, analizando el conjunto de pasos y el marco de trabajo, la distribución de las fases se presentan a continuación:

1. PARTE 1: Proceso de preparación de datos y de ingeniería de atributos

1.1. Fase 1

- Definir un problema de Machine Learning (ML) y proponer una solución

1.2. Fase 2

- Construir tu conjunto de datos

1.3. Fase 3

- Transformar los datos

1.4. Fase 4

- Entrenar el modelo

1.5. Fase 5

- Usar el modelo para hacer predicciones

Además, basándonos en el autor [35], las actividades consideradas para el marco de trabajo SCRUM, el cual sería la segunda parte de todo el desarrollo del sistema, son las siguientes:

2. PARTE 2: SCRUM

2.1. Inicio

- Alcance del proyecto
- Definiendo roles de trabajo
- Definiendo el producto backlog

2.2. Sprint 1 al 3

- Sprint planning
- Sprint backlog
- Historias de usuario
- Taskboard
- Daily scrum
- Sprint review
- Criterios de aceptación
- Resultado del sprint
- Sprint retrospective

Consideraciones éticas

Para garantizar la protección y bienestar, así como la seguridad de los datos de los participantes, en este caso para los miembros de la asociación, se han considerado que, durante la recopilación de datos a través de entrevistas y encuestas, se ha garantizado el anonimato de los participantes en las encuestas. La información almacenada dentro de la aplicación será confidencial y se utiliza solo para fines de investigación y pruebas del sistema. Los participantes y encargados de la asociación han sido informados sobre el propósito del trabajo de investigación y han otorgado su consentimiento para participar. Asimismo, existe la capacitación del personal involucrado en la gestión de la información para garantizar un manejo seguro de los datos a través de un Manual de usuario. Por otro lado, no se ha considerado la solicitud de ingreso de contraseñas debido a que se maneja autenticaciones anónimas dentro de la aplicación. A pesar de ello, se garantiza el resguardo de datos dentro del proveedor de almacenamiento como lo es Firebase para el registro únicamente de los nombres de las personas que desee agregar el usuario con

discapacidad visual. Es así que todos los datos recopilados se almacenan de manera segura y tienen acceso por usuario accedido en la aplicación. Se ha establecido ello para garantizar que solo las personas con la autorización adecuada puedan acceder a la información. Además, se mantiene un compromiso absoluto de mantener la confidencialidad de los datos y la información de los participantes en la investigación.

Resultados y discusión

En base a la metodología

Las metodologías utilizadas del proyecto fueron SCRUM y el Proceso de preparación de datos y de ingeniería de atributos.

Proceso de preparación de datos y de ingeniería de atributos

Los resultados complementarios se evidencian en el Anexo N°05 Proceso de preparación de datos y de ingeniería de atributos. Sin embargo, para llegar a la Fase 1, se tuvo que realizar una evaluación exhaustiva de los modelos de reconocimiento facial con la justificación de analizar la optimización de cada uno en cuanto a exactitud, nivel de profundidad, número de capas y de épocas. Los modelos de reconocimiento facial seleccionados para la evaluación fueron VggFace, Inceptionv3 y Facenet. Los resultados obtenidos por cada modelo en cuanto a accuracy se detallan en los resultados en base a los objetivos, específicamente en el segundo objetivo donde se estableció un cuadro comparativo con los indicadores y modelos ya mencionados.

Fase 1: Definir un problema de Machine Learning (ML) y proponer una solución

En este estudio se propuso una solución el cual se utilizó una arquitectura de red convolucional (CNN) que aprendió a mapear rostros en el espacio euclidiano. Una vez creado este espacio, la detección y verificación de rostros, así como la agrupación se pudieron implementar mediante algoritmos de clasificación que usan "embeddings" como feature vectors (vectores de características) [38]. A continuación, se presenta la arquitectura:

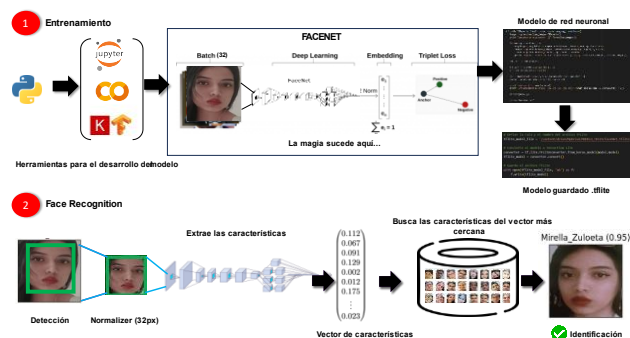


Fig. 1. Arquitectura del modelo de reconocimiento facial

Fase 2: Construir tu conjunto de datos

El conjunto constó de un total de 63 fotografías de 9 personas, 4 mujeres, 5 hombres, y una edad media de 23 años. El alojamiento del dataset fue en google drive [40].

Del conjunto se obtuvo:

- Promedio: 7 fotos por clase
- Desviación estándar: 2,21 fotos por clase
- Mínimo: 4 fotos por clase
- Máximo: 11 imágenes en una clase
- Total: 63 imágenes

Fase 3: Transformar los datos

Para la transformación de datos, fue necesario el preprocesamiento del conjunto de datos para el reconocimiento. Esto consistió en:

- Detectar rostros
- Alinear y recortar las caras de las fotos
- Cambiar el tamaño a 160x160 pixels

Los rostros de las imágenes fueron reconocidos mediante el Multi-Task Cascaded Convolutional Neural Networks (MTCNN) presentado por Paz [41]. El proceso descrito agregó "-N" al nombre de la foto, donde N fue la cantidad de rostros que fueron detectados en la fotos, por lo que, si hubo una foto con varios rostros, la imagen no fue sobrescrita.

Fase 4: Entrenar el modelo

Los parámetros del entrenamiento fueron:

- SVC hiper-parámetros:
 - o Kernel: lineal.
 - o Epochs: Número de iteraciones controlado por la implementación que terminó en la convergencia.
 - o C (penalización): 10.
- Cross validation: k-fold en 7.
- Conjunto de datos:
 - o 70% de fotos para entrenamiento.
 - o 30% de fotos para prueba.

Los valores del conjunto de datos se optaron por la relación de 70 a 30 para entrenamiento y prueba porque se usa a menudo en minería de datos [44].

Fase 5: Usar el modelo para hacer predicciones

Con respecto a las predicciones, considerando nuestro conjunto de datos, este devolvió una probabilidad por las 19 fotos de los datos de prueba, de las cuales se obtuvo:

- Media: 0.94
- Desviación estándar: 0.13
- Mínimo: 0.41
- Máximo: 1.00
- Total: 19 muestras

Finalizando esta secuencia de pasos, los resultados logrados en cuanto a precisión fueron de un 94% siendo un modelo adecuado para la posterior integración con la aplicación móvil.

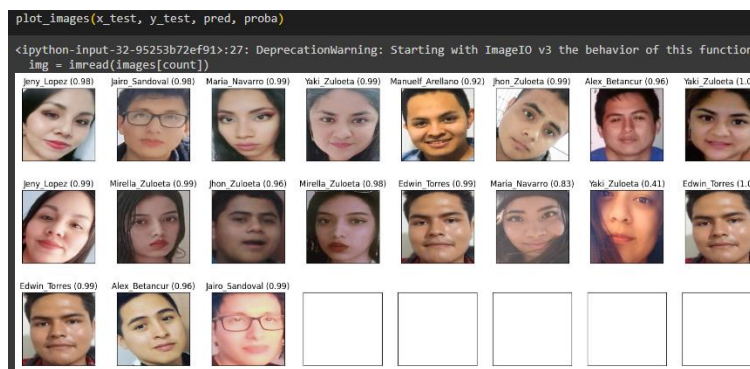


Fig. 2. Resultados de las predicciones del modelo entrenado

SCRUM

Inicio

En la presente tesis se desarrolló el sistema de visión computacional basada en reconocimiento facial para el apoyo a las personas con discapacidad visual de la organización regional de ciegos del Perú Chiclayo. A través del producto se visualizó cómo es que las personas que presentan discapacidad visual mejoraron su capacidad de independencia asegurando que la sociedad ya no los vea como personas que requieran de ayuda adicional. Asimismo, con la presente solución se les permitió a las personas reconocer a las demás personas que haya en su entorno. A continuación, se presentan los roles de Scrum:

TABLA III
ROLES DE SCRUM

Rol de trabajo	Nombres y Apellidos	Descripción
Product owner	Mery Ordinola Farias	Negocia decisiones sobre el producto acreditable.
Scrum máster	Oscar Alex Serquén Yparraguirre	Guía el equipo Scrum
Scrum team	Mirella Jazmín Zuloeta López	Encargada de todo el desarrollo de la aplicación.

Asimismo, se evidencia la arquitectura de todo el sistema de visión computacional:

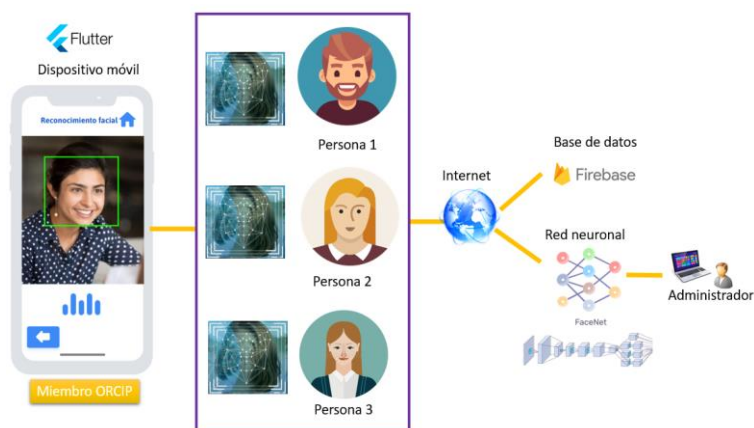


Fig. 3. Arquitectura del sistema de visión computacional

Sprint 1

Las historias de usuarios presentadas dentro del primer sprint backlog van enfocadas a la prioridad que demanda el trabajo de investigación lo cual, se puede apreciar en las interfaces de la Fig. 4.

TABLA IV
SPRINT BACKLOG DEL SPRINT 1

ID	Prioridad	Usuario	Enunciado	Sprint	Esfuerzo	Dimensión (días)
HU-01	M	Persona	Como persona, requiero tener pantallas amigables dentro de la aplicación que presenten un alto contraste en el diseño y no sean difíciles de utilizar	1	Medio	1
HU-02	M	Persona de ayuda	Como persona de ayuda, deseo agregar personas ya sea desde una foto de galería o una foto tomada instantánea a través de la cámara del dispositivo.	1	Alto	2
HU-03	M	Persona	Como persona, deseo reconocer a las personas agregadas de mi entorno en tiempo real mediante audio.	1	Alto	3

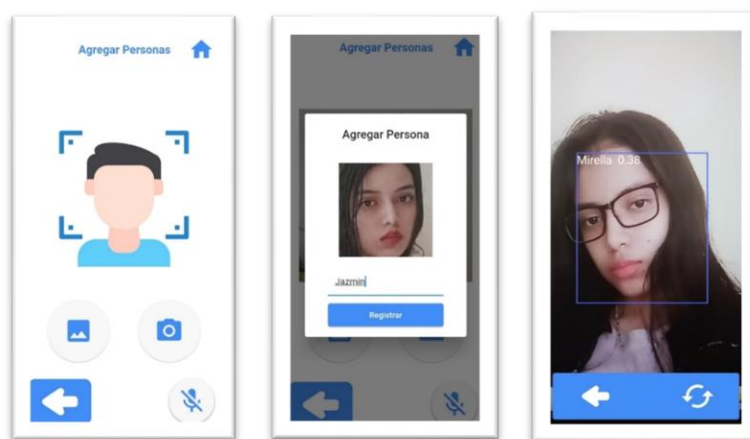


Fig. 4. Resultados del Sprint 1

Sprint 2

Con respecto al segundo sprint, en el sprint backlog se establece requerimientos donde fueron de igual forma consideradas como necesidad para las personas invidentes siendo incorporadas dentro de la presente tesis lo cual, se aprecia en las interfaces de la Fig. 5.

TABLA V
SPRINT BACKLOG DEL SPRINT 2

ID	Prioridad	Usuario	Enunciado	Sprint	Esfuerzo	Dimensión (días)
HU-04	M	Persona	Como persona, me gustaría contar con asistente de voz en toda la aplicación	2	Medio	2
HU-05	S	Persona	Como persona, deseo reconocer cualquier material bibliográfico a través de audio	2	Alto	3
HU-06	S	Persona de ayuda	Como persona de ayuda, deseo contar con un mantenimiento de todas las personas que se han registrado en la aplicación	2	Alto	2

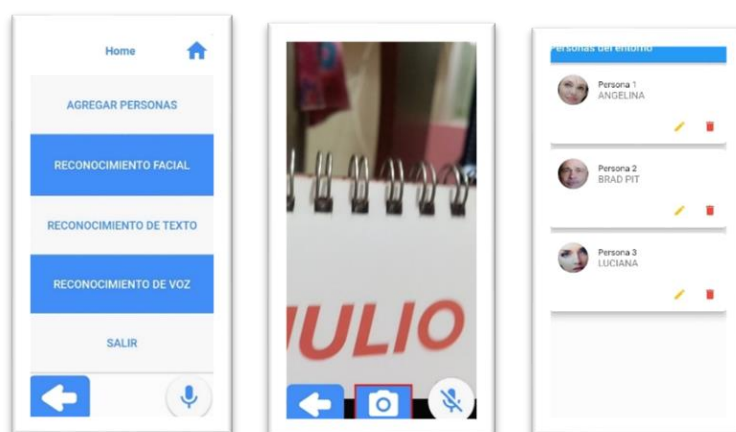


Fig. 5. Resultados del Sprint 2

Sprint 3

En este último sprint se elaboraron los requerimientos finales presentados en el sprint backlog, lo cual permitieron dar valor agregado para las personas con discapacidad visual para una mayor familiaridad con la aplicación. Los resultados se evidencian en las interfaces de la Fig. 6.

TABLA VI
SPRINT BACKLOG DEL SPRINT 3

ID	Prioridad	Usuario	Enunciado	Sprint	Esfuerzo	Dimensión (días)
HU-07	S	Persona	Como persona, me gustaría que también pueda identificar las emociones de las personas	3	Alto	4

HU-08	C	Persona	Como persona, deseo que me digan cuando toco algún componente de la aplicación como un asistente	3	Medio	3
-------	---	---------	--	---	-------	---

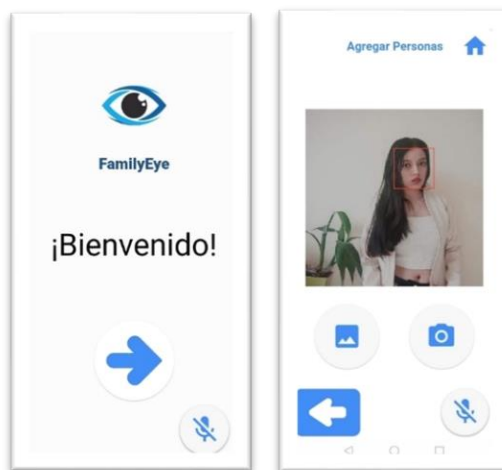


Fig. 6. Resultados del Sprint 3

En base a los objetivos de la investigación

Determinar la arquitectura del modelo que permita dar soporte al sistema de visión computacional para la ejecución del algoritmo.

De acuerdo con [48], se hizo la comparativa de arquitecturas de reconocimientos facial, los cuales son Facenet, GoogleNet y MobileNet. Con lo mencionado, se estableció un cuadro comparativo como medio de verificación en base a los indicadores de rendimiento, portabilidad y costos. Los valores ponderados para el cálculo son de 40%, 30% y 30% respectivamente de acuerdo a la importancia calificada por los autores [49]–[51], además de la naturaleza del trabajo de investigación donde se prioriza con mayor porcentaje al criterio de rendimiento debido a que correrá el modelo dentro de una aplicación móvil. La evaluación por cada arquitectura se ha considerado de acuerdo a la escala de Likert de 5 puntos [52], siendo calificado en base a las teorías que presenta el autor [48] ejecutándose así en la siguiente tabla:

TABLA VII
COMPARACIÓN DE ARQUITECTURAS EXISTENTES [48]

CRITERIO	FaceNet	GoogleNet	MobileNet
Rendimiento (40%)	5	3	5
Portabilidad (30%)	5	3	4
Costos (30%)	3	3	3

- **Para FaceNet:**

$$\text{Valor ponderado} = (5 * 0.4) + (5 * 0.3) + (3 * 0.3) = 2.0 + 1.5 + 0.9 = \mathbf{4.4}$$

- **Para GoogleNet:**

$$\text{Valor ponderado} = (3 * 0.4) + (3 * 0.3) + (3 * 0.3) = 1.2 + 0.9 + 0.9 = 3.0$$

- **Para MobileNet:**

$$\text{Valor ponderado} = (5 * 0.4) + (4 * 0.3) + (3 * 0.3) = 2.0 + 1.2 + 0.9 = 4.1$$

Considerando los resultados, la arquitectura idónea para un adecuado reconocimiento facial fue de Facenet, quedando posteriormente una arquitectura para todo el modelo propuesto visualizándose en la Fig. 1. Arquitectura del modelo de reconocimiento facial.

Construir el modelo de reconocimiento facial que se usará para el desarrollo del sistema de visión computacional.

Para corroborar el desarrollo de este objetivo, se hizo la evaluación y entrenamiento de los modelos existentes de reconocimiento facial para seleccionar el idóneo. Los modelos seleccionados fueron Facenet, VggFace e InceptionV3. Considerando ello, se obtuvieron diferentes resultados por cada uno contemplando los mismos indicadores. Todo ello se presenta en la siguiente tabla:

TABLA VIII
COMPARACIÓN DE MODELOS DE RECONOCIMIENTO FACIAL EXISTENTES

Modelos	Nº de capas	Nº de épocas	Nivel de profundidad del modelo	Accuracy del modelo
Facenet	3	12	512	98%
VggFace	6	46	4096	96.45%
Inceptionv3	3	256	2048	84%

Observando los resultados del accuracy de la Tabla XVII se pudo concluir que el modelo ideal de reconocimiento facial de “Facenet” con 98% de exactitud, alcanzando resultados satisfactorios y logrando ser accesible a las necesidades de la presente investigación. Los resultados de todo el entrenamiento posterior con un conjunto de datos personalizado se encuentran en el Anexo N° 05. Por otro lado, para corroborar finalmente la construcción del modelo, se hizo un checklist contrastando los indicadores de: Porcentaje de imágenes de entrenamiento, porcentaje de imágenes de prueba y el tiempo de ejecución del modelo.

TABLA IX
CONTRASTACIÓN DEL OBJETIVO EN BASE A INDICADORES

Indicador	Actividades	¿Realizado?	Descripción
Porcentaje de imágenes de entrenamiento	- Determinar el porcentaje de imágenes utilizado para el entrenamiento del modelo.	<input checked="" type="checkbox"/>	Se determinó el 70% de imágenes de entrenamiento del modelo.
	- Seleccionar y preparar el conjunto de datos de entrenamiento.	<input checked="" type="checkbox"/>	Se hizo uso del método <code>train_test_split</code> de <code>sklearn</code> para dividir en base al porcentaje establecido de 70% a las imágenes agrupadas de las personas por directorios a través de Google Drive.
	- Identificar las imágenes de entrenamiento con <code>data augmentation</code>	<input checked="" type="checkbox"/>	Después de haberse elaborado el procedimiento de <code>ata augmentation</code> , se identificó un total de 2584 imágenes aumentando en un 2550 imágenes gracias a esta técnica .
Porcentaje de imágenes de prueba	- Determinar el porcentaje de imágenes utilizado para evaluar el rendimiento del modelo.	<input checked="" type="checkbox"/>	Se determinó el 30% de imágenes de entrenamiento del modelo.
	- Preparar el conjunto de datos de prueba del modelo.	<input checked="" type="checkbox"/>	Se hizo uso del método <code>train_test_split</code> de <code>sklearn</code> para dividir en base al porcentaje establecido de 30% a las imágenes agrupadas de las personas por directorios a través de Google Drive.
Tiempo de ejecución del modelo	- Registrar el tiempo que tarda el modelo en entrenarse.	<input checked="" type="checkbox"/>	El tiempo empleado que tomó el modelo para su entrenamiento fue de 5 minutos con 50 segundos.
	- Medir el tiempo de ejecución para inferencias en nuevas imágenes de rostros.	<input checked="" type="checkbox"/>	El tiempo que tuvo el modelo para las inferencias antes nuevas imágenes fueron de 2 segundos.

Obtener una validación aprobatoria del modelo de reconocimiento facial a través de las principales métricas para visión computacional.

De acuerdo a las métricas de visión computacional que son exactitud, precisión y F1(Porcentaje de predicciones correctas generadas por la aplicación) se obtuvieron los siguientes resultados luego de haber sido evaluado el modelo con el conjunto de datos propio:

	precision	recall	f1-score	support
Alex_Betancur	1.00	1.00	1.00	2
Edwin_Torres	1.00	1.00	1.00	5
Jairo_Sandoval	1.00	1.00	1.00	4
Jeny_Lopez	0.75	1.00	0.86	3
Jhon_Zuloeta	1.00	1.00	1.00	3
ManuelF_Arellano	1.00	1.00	1.00	2
Maria_Navarro	1.00	1.00	1.00	2
Mirella_Zuloeta	1.00	1.00	1.00	3
Yaki_Zuloeta	1.00	0.80	0.89	5
accuracy			0.97	29
macro avg	0.97	0.98	0.97	29
weighted avg	0.97	0.97	0.97	29

Fig. 7. Resultados de la validación del modelo

Al observar los resultados, el modelo obtuvo un 97% de accuracy (exactitud) aproximadamente. Asimismo, considerando las siguientes métricas de precisión y F1 acorde a los resultados mostrados previamente, se puede observar que, en promedio, luego de haberse evaluado las predicciones por personas, se obtuvo un 98% y 97% respectivamente, lo cual se concluye que, de acuerdo a las métricas de visión computacional, sí se obtuvieron resultados mayores a un 70%.

Validar el funcionamiento del sistema de visión computacional con respecto a los criterios de usabilidad y utilidad percibida por los miembros de la Organización regional de ciegos del Perú – Chiclayo.

Para efectuar dicha validación del funcionamiento del sistema de visión computacional se tomó como referencia el cuestionario del Anexo N°04 que es el modelo de aceptación tecnológica (TAM). Al ser un caso de estudio, se consideró a la población de la asociación ORCIP Chiclayo lo cual se creyó conveniente el uso de un muestreo no probabilístico por juicio debido a la poca cantidad de miembros que se encuentran dentro de la asociación que son de 25 personas. Por ello, los criterios de selección fueron: Como criterio de inclusión se tuvo a las personas con ceguera total donde conforman de 13 personas; y como criterio de exclusión se tuvo al conjunto de las 12 personas restantes que presentan deficiencias visuales pero que de alguna u otra forma, logran visualizar el entorno. Es así que dicha cantidad de personas fue representativa para nuestra población evaluada. Con lo mencionado, la encuesta para dicho conjunto de personas está establecida en escala de Likert apreciándose de manera cualitativa en la Tabla XVIII. A razón de que las personas con discapacidad visual no pueden observar dicho cuestionario de manera legible, las preguntas fueron interrogadas de manera verbal quedando los resultados de la siguiente manera:

TABLA X
RESULTADOS DEL MODELO DE ACEPTACIÓN TECNOLÓGICA [53]

Facilidad de uso percibida (PEOU)	En absoluto	Poco	Neutro	Bastante	Completamente
PEOU1	-	-	-	-	100%
PEOU2	-	-	-	-	100%
PEOU3	-	-	-	25%	75%
PEOU4	-	-	-	25%	75%
PEOU5	-	-	-	45%	55%
Utilidad percibida (PU)	En absoluto	Poco	Neutro	Bastante	Completamente
PU1	-	-	-	-	100%
PU2	-	-	-	-	100%
PU3	-	-	-	-	100%
PU4	-	-	-	25%	75%
PU5	-	-	-	-	100%

De acuerdo a la tabla presentada anteriormente, el análisis de resultados se inicia primeramente con la facilidad de uso percibida (PEOU). Considerando la Fig. 14, los encuestados manifestaron lo siguiente: El 100% de las personas con discapacidad visual considera completamente a que la aplicación es fácil de usar (PEOU1). Asimismo, el 100% de los encuestados considera completamente a que el proceso de reconocimiento facial es sencillo de realizar (PEOU2).

Con relación a si la aplicación proporciona instrucciones claras y fáciles de seguir, el 75% de los encuestados considera completamente y el 25% considera bastante (PEOU3).

Sobre si la navegación de la aplicación es intuitiva el 75% considera completamente y el 25% considera bastante (PEOU4). Por último, acerca de si la aplicación presenta barreras de accesibilidad, el 55% considera completamente y el 45% considera bastante (PEOU5).

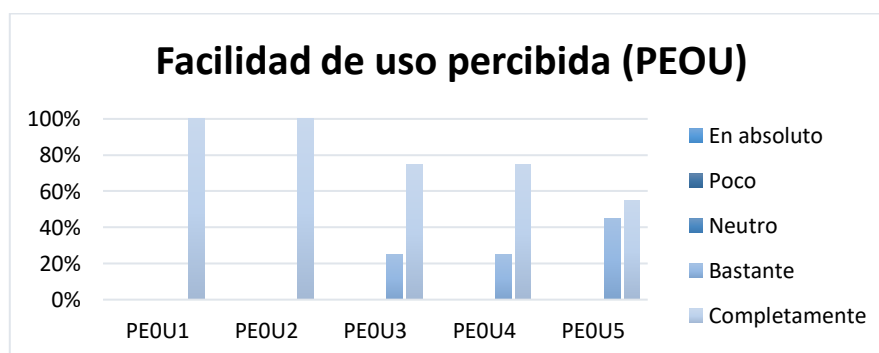


Fig. 8. Resultados de Facilidad de uso percibida por TAM

Por otra parte, continuando con el siguiente indicador que viene relacionado a la utilidad percibida (PU), los resultados de los encuestados considerando la Fig. 15

manifestaron lo siguiente: El 100% de las personas con discapacidad visual considera completamente a que la aplicación móvil de reconocimiento facial es útil para sus necesidades (PU1). Asimismo, el 100% de los encuestados considera completamente a que sienten gran efectividad del reconocimiento facial en la aplicación (PU2). Del mismo modo, el 100% de los encuestados considera completamente a que la aplicación sí proporciona una experiencia satisfactoria en cuanto a la identificación de personas (PU3).

Acerca de si la aplicación cumple con las expectativas en cuanto a la precisión del reconocimiento facial, el 75% de las personas consideraron completamente y el 25% restante consideró bastante (PU4). Por último, el 100% de los encuestados creyó que el sistema de reconocimiento facial mejora su independencia en la identificación de personas (PU5).

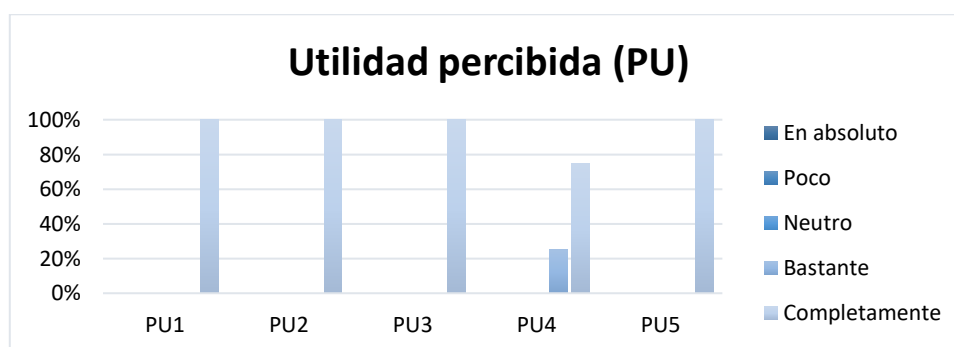


Fig. 9. Resultados de Utilidad percibida por TAM

Finalmente, en base a las Fig.14 y Fig. 15 se identificó que en cuanto a facilidad de uso percibida y utilidad de uso, los encuestados estuvieron inclinados a la categoría de Completamente, equivalente a una puntuación de 5 en escala de Likert, lo cual para PEOU se obtuvo un 81% y en PU se obtuvo un 95% en promedio de la más alta categoría, siendo los resultados satisfactorios para este trabajo de investigación.

Discusión

De acuerdo con el primer antecedente internacional de Yung [11], considerando sus resultados obtenidos en su informe, se pudo evidenciar que hizo uso de un algoritmo de reconocimiento de objetos como core principal, además de funcionalidades idóneas adicionales a diferencia de la presente tesis. Del mismo modo, el autor optó por tomar dicho modelo con datos ya entrenados a diferencia de este presente informe que se optó por realizar un conjunto de datos personalizados. A pesar de ello, con respecto a funcionalidades de asistencia y la esencia del diseño de interfaces, estas fueron tomadas

en cuenta para la incorporación de la presente tesis dirigida a personas con discapacidad visual.

Del mismo modo, en el estudio realizado por Golberg [12], se analizaron sus resultados obtenidos, y en su propuesta consideró interfaces intuitivas para una aplicación móvil capaz de ser sencilla para las personas con discapacidad visual. A pesar de que sus funcionalidades no son muy relacionadas a comparación de la investigación de [11], esta nos permitió tener un mayor análisis de qué es lo principal que se debe incorporar para la implementación de una herramienta de asistencia para el público objetivo como herramientas de audio y una navegación en dos pasos considerando a las complicaciones que podría tener el público objetivo cuando hace uso de una aplicación móvil.

Asimismo, considerando el trabajo de Márquez et al. [13], se presenta un modelo híbrido de memorias asociativas la cual fue abordado para personas invidentes. Este enfoque incorpora técnicas interesantes de reconocimiento facial asemejándose a la estrategia aplicada a la aplicación móvil de esta tesis. Sin embargo, es relevante destacar que en esta investigación se optó por la evaluación de algoritmos de reconocimiento facial con arquitecturas robustas de redes neuronales convolucionales (CNN) en lugar del modelo híbrido de memorias asociativas. Además, el modelo se entrenó con un conjunto de datos accesibles en lugar de datasets abiertos, teniendo como precisión un 99% lo que hizo que el modelo fuera más propio.

En cuanto a Ríos y Guerrero [16], los autores nos brindaron el uso de las técnicas de inteligencia artificial lo cual fue abordado eficazmente en su investigación de implementación de un sistema inteligentes generador de ordenes de desplazamiento. Sin embargo, es importante destacar que en esta investigación se optó por el uso de reconocimiento facial a través de dispositivos móviles en lugar del uso de un software que no es portable para las personas con discapacidad visual. A pesar de que el uso de las herramientas tecnológicas es notorio debido a que presenta diferentes soluciones, se abordaron claramente en comparativa de su investigación con la presente tesis, el apoyo a las personas invidentes haciendo uso de reconocimiento facial.

Finalmente, considerando a Sánchez [17], se toma como referencia igualmente el uso de soluciones móviles como en [11], [12], [14] sobre el uso de los principios de usabilidad específicamente para personas con discapacidad visual. En el enfoque del autor incorpora reconocimiento de palabras además de herramientas para el sistema braille. No obstante, es conveniente precisar que en esta investigación se tomó por el reconocimiento

de rostros mediante algoritmos robustos en lugar de reconocimiento de palabras que se usa otro tipo de técnicas para la clasificación de imágenes.

Conclusiones

Con la presente tesis se puede llegar como primera conclusión que a través de una herramienta de visión computacional basada en reconocimiento facial se puede apoyar a los miembros de la Organización regional de ciegos del Perú – Chiclayo para que puedan identificar personas en un entorno.

Para determinar la arquitectura del modelo que permita dar soporte al sistema de visión computacional para la ejecución del algoritmo, se llevó a cabo la comparativa de arquitecturas existentes y en base a los resultados, se concluye que Facenet fue la arquitectura idónea debido a que alcanzó un valor de 4.4 de ponderado.

Para construir el modelo de reconocimiento facial para el desarrollo del sistema de visión computacional, se hizo una evaluación de los modelos existentes de reconocimiento facial y definiendo el más robusto en lo que respecta a la solución, se concluye que Facenet obtuvo una alta precisión con un 98%, siendo así el modelo ideal posteriormente efectuado para el entrenamiento con un conjunto de datos personalizado.

Para obtener una validación aprobatoria del modelo de reconocimiento facial, se llevaron a cabo pruebas a través de las principales métricas de visión computacional. En cuanto a exactitud, precisión y F1 se obtuvo un 98%, 97% y 97% respectivamente, por lo que se concluye que, en base a los resultados brindados, el modelo alcanzó resultados superables a un 70%.

Con el objetivo de validar el funcionamiento del sistema de visión computacional con respecto a los criterios de usabilidad y utilidad percibida por los miembros de la Organización regional de ciegos del Perú – Chiclayo, se realizó a través del cuestionario de aceptación tecnológica dirigida a las personas con discapacidad visual por lo que se concluye que el trabajo es aceptable dado que se obtuvieron como resultados un 81% de facilidad de uso y 95% de utilidad percibida.

Recomendaciones

A continuación, se establecen las recomendaciones del presente trabajo de tesis:

1. Se recomienda que la aplicación no haga solo el reconocimiento a través de fotos, sino también basada en los movimientos de las personas mediante un Convolutional Recurrent Neural Network (CRNN).


2. Gracias a que contamos con dispositivos móviles para que se pueda realizar el reconocimiento, se recomienda que en un futuro se pueda integrar la detección a través de una cámara incorporada en los lentes de las personas con discapacidad visual para que más adelante no puedan depender totalmente de un dispositivo móvil.
3. Finalmente se recomienda que la aplicación funcione como un servicio dentro del dispositivo móvil como Talkback (Android) o VoiceOver (iOS) para que el usuario acceda directamente ni bien encienda su celular.

Referencias

- [1] «Ceguera y discapacidad visual». Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment>
- [2] M. Mukhiddinov y J. Cho, «Smart Glass System Using Deep Learning for the Blind and Visually Impaired», *Electronics*, vol. 10, n.º 22, Art. n.º 22, ene. 2021, doi: 10.3390/electronics10222756.
- [3] U. Sait *et al.*, «Design and development of an assistive device for the visually impaired», *Procedia Computer Science*, vol. 167, pp. 2244-2252, 2020, doi: 10.1016/j.procs.2020.03.277.
- [4] «UNICEF: “Existen casi 240 millones de niños con discapacidad en todo el mundo; en RD el 10% de niños entre 2 y 17 años tiene algún tipo de discapacidad motora, auditiva o visual”». Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://www.unicef.org/dominicanrepublic/comunicados-prensa/unicef-existen-casi-240-millones-de-ninos-con-discapacidad-en-todo-el-mundo>
- [5] «MonográficoFOAL_0.pdf». Accedido: 8 de octubre de 2023. [En línea]. Disponible en: https://www.foal.es/sites/default/files/Monogr%C3%A1ficoFOAL_0.pdf
- [6] «Instituto Nacional de Estadística e Informática». Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://m.inei.gob.pe/prensa/noticias/en-el-pais-existen-3-millones-209-mil-261-personas-con-discapacidad-11687/>
- [7] «Informe-CONADIS.pdf». Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://www.conadisperu.gob.pe/observatorio/wp-content/uploads/2021/05/Informe-CONADIS.pdf>
- [8] «Organización Regional De Ciegos Del Perú | FOAL». Accedido: 8 de octubre de 2023. [En línea]. Disponible en:

<https://www.foal.es/es/asociaciones/organizaci%C3%B3n-regional-de-ciegos-del-per%C3%BA>

- [9] D. Lee y J. Cho, «Automatic Object Detection Algorithm-Based Braille Image Generation System for the Recognition of Real-Life Obstacles for Visually Impaired People», *Sensors*, vol. 22, n.º 4, Art. n.º 4, ene. 2022, doi: 10.3390/s22041601.
- [10] J. D. Watters, A. Hill, M. Weinrich, C. Supalo, y F. Jiang, «An Artificial Intelligence Tool for Accessible Science Education», *jSESD*, vol. 24, n.º 1, pp. 1-14, sep. 2021, doi: 10.14448/jsesd.13.0010.
- [11] D. T. Y. Shen, «Mobile Application for The Visually Impaired».
- [12] M. Golberg, «A Mobile Cyber-Physical System Framework for Aiding People with Visual Impairment», Ph.D. dissertation, CUNY Univ., New York, N.Y., 2019.
- [13] M. Márquez-Olivera, A.-G. Juárez-Gracia, V. Hernández-Herrera, A.-J. Argüelles-Cruz, y I. López-Yáñez, «System for Face Recognition under Different Facial Expressions Using a New Associative Hybrid Model Am α -KNN for People with Visual Impairment or Prosopagnosia», *Sensors*, vol. 19, n.º 3, Art. n.º 3, ene. 2019, doi: 10.3390/s19030578.
- [14] C. Araujo, C. J. Eduardo, A. Cajas, y Y. Roy, «TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE: INGENIERO DE SISTEMAS».
- [15] H. R. Herencia Aliaga y D. O. Heredia Leon, «Solución tecnológica para la inclusión de alumnos invidentes en universidades», Pregrado, Universidad Peruana de Ciencias Aplicadas (UPC), Perú, 2020. doi: 10.19083/tesis/653457.
- [16] «Ríos Kavadoy Juan Carlos - Guerrero Bello Harold Gianpierre.pdf». Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://repositorio.upn.edu.pe/bitstream/handle/11537/23430/R%C3%ADos%20Kavadoy%20Juan%20Carlos%20-%20Guerrero%20Bello%20Harold%20Gianpierre.pdf?sequence=8&isAllowed=y>
- [17] E. L. S. Tapia, «Aplicativo para el reconocimiento de palabras basado en un dispositivo gráfico táctil en el aprendizaje con el sistema Braille, en el centro de recursos de educación básica especial “Crebe” de Chiclayo».
- [18] R. J. F. Llontop, «BASTÓN SENSORIAL GEOLOCALIZADOR INTELIGENTE PARA APOYAR EN EL DESPLAZAMIENTO DE PERSONAS INVIDENTES EN LA ORGANIZACIÓN REGIONAL DE CIEGOS DEL PERÚ – CHICLAYO».
- [19] L. A. Z. Bispo, «INGENIERO DE SISTEMAS Y COMPUTACIÓN».

- [20] J. Arias-Urbe, Y. Llano-Naranjo, E. Astudillo-Valverde, y J. C. Suárez-Escudero, «Caracterización clínica y etiología de baja visión y ceguera en una población adulta con discapacidad visual», *RMO*, vol. 92, n.º 4, p. 989, oct. 2020, doi: 10.24875/RMO.M18000033.
- [21] D. L. Demmin y S. M. Silverstein, «Visual Impairment and Mental Health: Unmet Needs and Treatment Options», *Clinical Ophthalmology*, vol. 14, pp. 4229-4251, dic. 2020, doi: 10.2147/OPHTH.S258783.
- [22] P. Zamora López y C. Marín Perabá, «Tiflotecnologías para el alumnado con discapacidad visual», *ACADEMO (Asunción)*, vol. 8, n.º 1, pp. 109-118, feb. 2021, doi: 10.30545/academo.2021.ene-jun.10.
- [23] E. Soria, P. Rodríguez, Q. García, F. Vaquer, J. Vicent, y J. Vila, *Inteligencia Artificial*. Ra-Ma Editorial, 2022.
- [24] T. D. Mínguez, *Visión artificial: Aplicaciones prácticas con OpenCV - Python*. Marcombo, 2021.
- [25] T. Nasir, M. K. Malik, y K. Shahzad, «MMU-OCR-21: Towards End-to-End Urdu Text Recognition Using Deep Learning», *IEEE Access*, vol. 9, pp. 124945-124962, 2021, doi: 10.1109/ACCESS.2021.3110787.
- [26] A. P. y Madrid, *El reconocimiento facial es un superpoder. Cómo te afectará y por qué deberías conocerlo*. Dykinson, 2021.
- [27] F. Sabry, *Biometría: El futuro representado en la película «Minority Report» ya está aquí*. One Billion Knowledgeable, 2022.
- [28] V. M. García, *Creación de aplicaciones con Android*. Ra-Ma Editorial, 2021.
- [29] M. Barrientos Rodríguez *et al.*, «Metodología para el desarrollo del sistema Web para la gestión de los programas de maestría del Instituto “Pedro Kourí”», *Revista Cubana de Informática Médica*, vol. 14, n.º 2, dic. 2022, Accedido: 15 de octubre de 2023. [En línea]. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S1684-18592022000200008&lng=es&nrm=iso&tlng=es
- [30] R. McCarthy, *Agile y Scrum*. Amazon Digital Services LLC - Kdp, 2020. Accedido: 15 de octubre de 2023. [En línea]. Disponible en: https://www.google.com.pe/books/edition/Agile_y_Scrum/j9wezgEACAAJ?hl=es-419
- [31] «Las 5 fases de la metodología Scrum para tus proyectos ». Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://alvaroarrarte.com/fases-de-la-metodologia-scrum/>

- [32] «Preparación de datos e ingeniería de atributos en AA | Machine Learning», Google for Developers. Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://developers.google.com/machine-learning/data-prep?hl=es-419>
- [33] «El proceso de preparación de datos y de ingeniería de atributos | Machine Learning», Google for Developers. Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://developers.google.com/machine-learning/data-prep/process?hl=es-419>
- [34] «Manual de Frascati 2015: Guía para la recopilación y presentación de información sobre la investigación y el desarrollo experimental | es | OCDE | OECD». Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <https://www.oecd.org/publications/manual-de-frascati-2015-9789264310681-es.htm>
- [35] L. A. Aguilar Capcha, «Desarrollo e implementación de un sistema web basado en la metodología scrum, para mejorar el proceso de autorizaciones e infracciones en la dirección regional de transportes y comunicaciones de Ayacucho», *Development and implementation of a web system based on the scrum methodology, to improve the process of authorizations and infractions in the directorate of transport and communications of Ayacucho*, jul. 2021, Accedido: 8 de octubre de 2023. [En línea]. Disponible en: <http://repositorio.autonoma.edu.pe/handle/20.500.13067/1318>
- [36] Y. Taigman, M. Yang, M. Ranzato, y L. Wolf, «DeepFace: Closing the Gap to Human-Level Performance in Face Verification», en *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA: IEEE, jun. 2014, pp. 1701-1708. doi: 10.1109/CVPR.2014.220.
- [37] D. Sandberg, «Face Recognition using Tensorflow». 22 de octubre de 2023. Accedido: 22 de octubre de 2023. [En línea]. Disponible en: <https://github.com/davidsandberg/facenet>
- [38] S. Hettiarachchi, «Analysis of different face detection and recognition models for Android», M.S. tesis, M.S Univ., Sweden, SE, 2021.
- [39] «Labelled Faces in the Wild (LFW) Dataset». Accedido: 8 de junio de 2023. [En línea]. Disponible en: <https://www.kaggle.com/datasets/jessicali9530/lfw-dataset>
- [40] J. Elias, A. Rojas, L. Rojas, M. Alfaro, R. Uscamayta, y A. Gamarra, «Aplicación del deep learning para el reconocimiento facial con la presencia de oclusiones en el contexto de la pandemia covid 2021», *Rev. ECIPeru*, vol. 18, n.º 1, pp. 17-24, 2021, doi: 10.33017/RevECIPeru2021.0002/.

- [41] I. de P. Centeno, «ipazc/mtcnn». 8 de junio de 2023. Accedido: 8 de junio de 2023. [En línea]. Disponible en: <https://github.com/ipazc/mtcnn>
- [42] A. Rajagopalan, «Real-Time Deep Learning-Based Face Recognition System».
- [43] «scikit-learn: machine learning in Python — scikit-learn 1.2.2 documentation». Accedido: 8 de junio de 2023. [En línea]. Disponible en: <https://scikit-learn.org/stable/>
- [44] A. De, «LÍNEA DE INVESTIGACIÓN».
- [45] B. Barrera y J. Carlos, «Desarrollo de una aplicación híbrida inteligente y un modelo de análisis de datos en fases para resumir y representar los comentarios de clientes sobre productos y servicios de bares y restaurantes mediante Django, IONIC, Angular, Deep Learning y redes sociales».
- [46] L. A. R. Pino, «Memoria para optar al título de Ingeniero Civil en Informática.».
- [47] «Arraya et al. - PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO CIVI.pdf». Accedido: 15 de octubre de 2023. [En línea]. Disponible en: https://repositorio.urp.edu.pe/bitstream/handle/20.500.14138/3849/T030_45355923_T%20%20MEDINA%20LIMA%20LUIS%20ENRIQUE.pdf?sequence=1

Anexos**ANEXO N° 01. CONSTANCIA DE APROBACIÓN DEL PRODUCTO
ACREDITABLE DE LA ENTIDAD DONDE SE EJECUTÓ LA TESIS****CONSTANCIA DE APROBACIÓN DEL PRODUCTO ACREDITABLE****CARTA N°001-2023**

Estimado,

Ing. Mera Montenegro Huilder Juanito

Por medio de la presente certifico la aprobación de la tesis titulada "SISTEMA DE VISIÓN COMPUTACIONAL BASADA EN RECONOCIMIENTO FACIAL PARA IDENTIFICACIÓN DE PERSONAS DIRIGIDO A LOS MIEMBROS DE LA ORGANIZACIÓN REGIONAL DE CIEGOS DEL PERÚ – CHICLAYO" de la estudiante **ZULOETA LOPEZ MIRELLA JAZMIN**, perteneciente a la escuela profesional de **INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**; como producto acreditable de su trabajo de investigación de fin grado, cumpliendo los requisitos establecidos y objetivos planteados por este. El producto acreditable ha sido ejecutado y evaluado de acuerdo a los lineamientos y criterios establecidos por la asociación. Asimismo, la estudiante ha demostrado un compromiso excepcional y capacidad para llevar a cabo una investigación de alta calidad. Se expide la presente constancia para los fines que estime conveniente.

Atentamente, Sra. Mery Ordinola Farias.

Sra. Mery Ordinola Farias

**Chiclayo, 25 de octubre del 2023**

**ANEXO N° 02. ENTREVISTA AL REPRESENTANTE DE LA ORGANIZACIÓN
REGIONAL DE CIEGOS DEL PERÚ-CHICLAYO**

Objetivo: Identificar mediante cifras estadísticas acerca de la problemática actual que presentan los miembros de la asociación.

1. ¿Cuál es la cantidad de personas acogidas en la asociación ORCIP-Chiclayo?
2. ¿Qué problemas atraviesan los miembros de la asociación ORCIP-Chiclayo?
3. ¿Cuál es el porcentaje de los miembros de la asociación que presentan problemas de ceguera parcial?
4. ¿Cuál es el porcentaje de los miembros de la asociación que presentan problemas de ceguera total?
5. ¿Cuáles fueron las causas para que las personas se encuentren acogidas dentro de la asociación ORCIP-Chiclayo?
6. ¿Qué grado de dificultades presentan los miembros de la asociación ORCIP-Chiclayo?
7. ¿Qué tipos de accidentes han sufrido las personas con discapacidad visual?
8. ¿Cuál es el tipo de discriminación que han sufrido las personas con discapacidad visual de ORCIP-Chiclayo?
9. ¿Cómo ve la sociedad a los miembros de la asociación ORCIP-Chiclayo?
10. ¿Qué proyectos tecnológicos se han implementado para solucionar problemas a los miembros de la asociación ORCIP-Chiclayo durante estos últimos años?

**ANEXO N° 03. ENTREVISTA A UNA PERSONA CON DISCAPACIDAD VISUAL
DE LA ORGANIZACIÓN REGIONAL DE CIEGOS DEL PERÚ-CHICLAYO**

Objetivo: Identificar las dificultades que presentan los miembros de la asociación.

1. ¿Qué problemas presenta diariamente en la sociedad?
2. ¿Logra identificar a los miembros de su familia cuando se están desplazando de un lugar a otro?
3. ¿Logra identificar a todas las personas que están dentro de tu alrededor?
4. Si ustedes están en una reunión grande, ¿Lograría identificar qué persona le está hablando que no fuera su familiar?
5. ¿Qué problemas atraviesa cuando no logran identificar a alguien de su alrededor?
6. ¿De qué manera usa su dispositivo móvil?
7. ¿Ha escuchado de algunas tecnologías que permiten ayudarle durante la vida cotidiana?
8. ¿Cuáles serían las tecnologías ideales que le gustaría usar a diario?
9. ¿Les gustaría usar una solución tecnológica capaz de identificar qué persona está cerca de su entorno?
10. ¿Qué características le gustaría que presente dicha solución para que pueda identificar personas y pueda usarlo de manera sencilla?

ANEXO N° 04. MODELO DE ACEPTACIÓN TECNOLÓGICA

Tabla 1: Modelo de aceptación tecnológica (TAM)

DIMENSIONES SOBRE LA APLICACIÓN	ÍTEM	PREGUNTA	EN ABSOLUTO	POCO	NEUTRO	BASTANTE	COMPLETAMENTE
Facilidad percibida de uso (PEOU)	PEOU1	¿Considera que la aplicación es fácil de usar?	1	2	3	4	5
	PEOU2	¿Encuentra que el proceso de reconocimiento facial es sencillo de realizar?	1	2	3	4	5
	PEOU3	¿La aplicación proporciona instrucciones claras y fáciles de seguir?	1	2	3	4	5
	PEOU4	¿La navegación en la aplicación es intuitiva?	1	2	3	4	5
	PEOU5	¿La aplicación no presenta barreras de accesibilidad?	1	2	3	4	5
Utilidad percibida (PU)	PU1	¿Cree que la aplicación móvil de reconocimiento facial es útil para sus necesidades?	1	2	3	4	5
	PU2	¿Siente efectividad del reconocimiento facial en la aplicación?	1	2	3	4	5
	PU3	¿La aplicación proporciona una experiencia satisfactoria en la identificación de personas?	1	2	3	4	5
	PU4	¿La aplicación cumple con sus expectativas en cuanto a la precisión del reconocimiento facial?	1	2	3	4	5
	PU5	¿Cree que el sistema de reconocimiento facial mejora su independencia y seguridad en la identificación de personas?	1	2	3	4	5

ANEXO N° 05. PROCESO DE PREPARACIÓN DE DATOS E INGENIERÍA DE ATRIBUTOS

- Fase 2: Construir tu conjunto de datos

```
l = np.array(list(Counter(dataset_y).keys()))
v = np.array(list(Counter(dataset_y).values()))

'Mean: {:.2f}, Std: {:.2f}, Min: {:.2f}, Max: {:.2f}, Total: {}'.format(v.mean(), v.std(),
                                                                    v.min(), v.max(), v.sum())
'Mean: 7.00, Std: 2.21, Min: 4.00, Max: 11.00, Total: 63'
```

Ilustración 1: Medidas estadísticas de los datos

- Fase 3: Transformar los datos

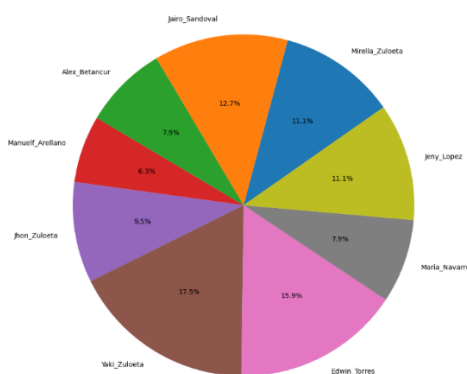


Ilustración 2: Distribución de los datos por cada clase

```
def preprocess_image(image_path, image_output_dir, align, crop_dim, margin):
    identity = os.path.basename(image_path).split('.')[0]
    output_path = os.path.join(image_output_dir, identity)

    logger.info('Leyendo la imagen: {}'.format(image_path))
    image = cv2.imread(image_path)
    image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)

    if len(glob('{}*.jpg'.format(output_path))) < 1:
        faces = detect_faces(image, margin)
        logger.info('Faces: {}'.format(len(faces)))

        for index, face in enumerate(faces):
            if align:
                image = align_face(image, face['keypoints'])
            cropped = image[face['right']:face['top'], face['left']:face['bottom'], :]
            cropped = cv2.resize(cropped, crop_dim, interpolation = cv2.INTER_AREA)
            cropped = cv2.cvtColor(cropped, cv2.COLOR_BGR2RGB)
            # Store cropped
            cv2.imwrite('{}-{}.jpg'.format(output_path, index), cropped)
```

Ilustración 3: Lógica del preprocesamiento de las imágenes

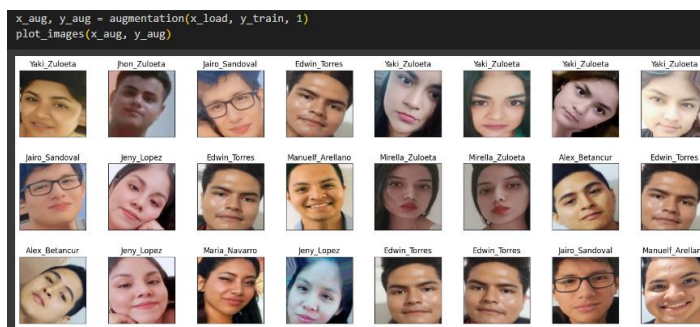


Ilustración 4: Desplazamiento de los rostros aplicando Data Augmentation

- Fase 4: Entrenar el modelo

```
lencoder, clf = train(x_train, y_train, c=10, cv=3, num_aug_iterations=150)
<ipython-input-31-c8bfcab0126>:6: DeprecationWarning: Starting with ImageIO v3 the behavior of this function
img = imread(filepath)
Images to process: 44
Images after augmentation: 3344 (+3300)
Scores: [1. 1. 1.]
Train accuracy: 1.00 (+/- 0.00)
```

Ilustración 5: Resultados del entrenamiento del modelo

- Fase 5: Usar el modelo para hacer predicciones

```
#PROBAR EL MODELO
embs, pred, proba = infer(lencoder, clf, x_test)
<ipython-input-31-c8bfcab0126>:6: DeprecationWarning: Starting with ImageIO v3 the behavior of this function
img = imread(filepath)
<
y_pred = lencoder.transform(pred)
y_true = lencoder.transform(y_test)
'Test accuracy: {:.3f}'.format(metrics.accuracy_score(y_true, y_pred))
'Test accuracy: 1.000'
```

Ilustración 6: Resultados del funcionamiento de la predicción del modelo

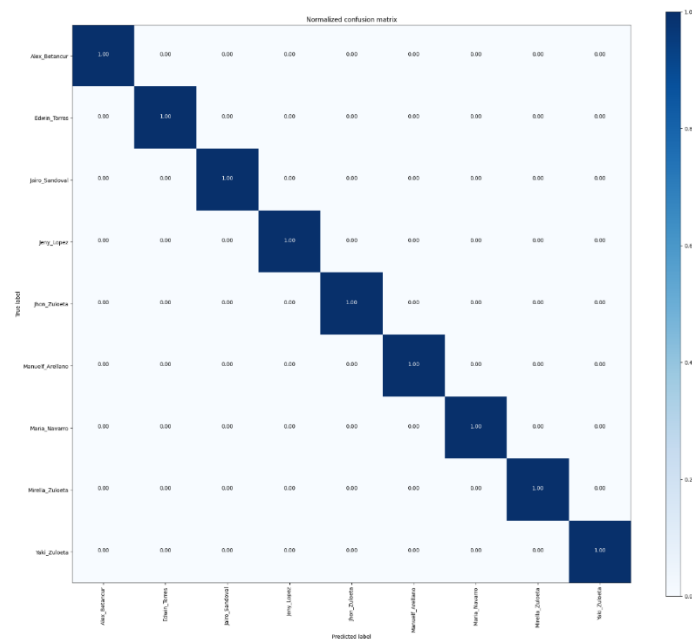


Ilustración 7: Matriz de confusión de la data de prueba distribuida en clase con las predicciones