



UNIVERSIDAD CÉSAR VALLEJO

FACULTAD DE INGENIERÍA Y ARQUITECTURA

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

**Implementación de un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes y video de los trabajadores de la empresa Servicios Generales y Transporte
“Fresco Pez” Chulucanas – 2021**

AUTOR(ES):

GALLO VALLADARES, Cristhian Valentin (ORCID: 0000-0003-2847-3478)

NUNURA GARCIA, Diana Estefany (ORCID: 0000-0003-4955-6096)

ASESOR:

Ing. MORE VALENCIA, Rubén Alexander (ORCID: 0000-0002-7496-3702)

LÍNEA DE INVESTIGACIÓN:

Sistema de información y comunicaciones.

PIURA — PERÚ

2021

Dedicatoria

Esta investigación se la dedicamos a nuestros padres, y todos nuestros familiares que nos apoyaron a lo largo de nuestra carrera universitaria, a nuestro asesor y maestros por brindarnos su apoyo, enseñanza y amistad en estos años universitarios.

Agradecimiento

Agradecemos a nuestros padres y familiares por el apoyo afectivo, moral y económico brindado, a mis docentes, amigos y al gerente de la empresa Servicios Generales y Transporte Fresco Pez S.R.L. quien nos dio su apoyo y facilidades para realizar esta investigación. Le agradecemos a nuestro asesor, el ing. Rubén More Valencia por la confianza y apoyo a lo largo de todo el proceso de desarrollo de nuestra investigación.

Índice de contenidos

Resumen	7
Abstract	8
I. INTRODUCCIÓN	9
II. MARCO TEÓRICO.	12
III. METODOLOGÍA	18
3.1. Tipo y diseño de investigación.....	18
3.2. Variables y operacionalización	19
3.3. Población, muestra y muestreo	20
3.4. Técnicas e instrumentos de recolección de datos	21
3.5. Procedimientos	21
3.6. Método de análisis de datos	22
3.7. Aspectos éticos.....	23
IV. RESULTADOS	24
V. DISCUSIÓN	32
VI. CONCLUSIONES	36
VII. RECOMENDACIONES	37
REFERENCIAS	38
ANEXOS.	41

Índice de tablas

Tabla 1. Matriz de operacionalización de la variable.	19
Tabla 2. Comprobación del modelo principal.....	28
Tabla 3. Comprobación del modelo secundario.	30
Tabla 4. Tabla de frecuencias.	31
Tabla 5. Ficha de registro de la dimensión fondo de imagen.	41
Tabla 6. Ficha de registro de la dimensión medición matricial.	42
Tabla 7. Ficha de registro de la dimensión calidad de imagen.	43
Tabla 8. Ficha de registro de la dimensión resolución de imagen.....	44
Tabla 9. Operacionalización de variable.....	45

Índice de figuras

Figura 1. Ficha de registro N°1.....	24
Figura 2. Ficha de registro N°2.....	25
Figura 3. Ficha de registro N°3.....	26
Figura 4. Ficha de registro N°4.....	27
Figura 5. Imágenes con resultado erróneo.....	29
Anexo 1. Consentimiento informado.....	47
Anexo 2. Código para obtener los fotogramas.	48
Anexo 3. Directorio de imágenes.....	48
Anexo 4. Código para entrenar el modelo.	49
Anexo 5. Código para reconocer gestos.....	50

Resumen

El procesamiento digital de imágenes no es un tema tratado con normalidad en esta parte del país, pero se sabe que todos los días los seres humanos realizan inconscientemente dicho proceso, todo eso gracias a uno de los sentidos, la vista. Observar una escena y saber lo que se está realizando en esa escena es gracias a varios factores, como por ejemplo el uso de herramientas o insumos, esto se puede hacer para situaciones sencillas. Por otra parte, en las situaciones complejas, la observación humana no es totalmente efectiva, es por ello que en esta investigación se plantea la implementación de un sistema para el reconocimiento de gestos corporales usando lo antes mencionado, el procesamiento de imágenes. Este sistema busca ser implementado en la empresa Servicios Generales y Transporte Fresco Pez S.R.L. de la ciudad de Chulucanas. Para ello se han tomado las medidas necesarias para tener acceso a los materiales necesarios, en este caso lo más importante son los archivos de video del sistema de seguridad del establecimiento. La investigación empieza con la revisión de los antecedentes que han estudiado el tema propuesto, encontrando variedad de ellos y en los idiomas inglés y español. Como variable de estudio se definió el procesamiento de imágenes, desglosándola en dimensiones e indicadores que ayudan a la operacionalización de misma. Además, se han tomado en cuenta las técnicas e instrumentos con el fin de recopilar los datos necesarios para el ordenamiento de la información, otro punto importante en el desarrollo de la investigación son los aspectos éticos y administrativos.

Palabras clave: imagen digital, procesamiento de imagen de video, pixeles de video, gestos corporales, redes neuronales convolucionales.

Abstract

Digital image processing is not a subject treated normally in this part of the country, but it is known that every day human beings unconsciously perform this process, all thanks to one of the senses, sight. Observing a scene and knowing what is being performed in that scene is thanks to several factors, such as the use of tools or inputs, this can be done for simple situations. On the other hand, in complex situations, human observation is not fully effective, that is why in this research we propose the implementation of a system for body gesture recognition using the above mentioned, image processing. This system seeks to be implemented in the company Servicios Generales y Transporte Fresco Pez S.R.L. in the city of Chulucanas. For this purpose, the necessary measures have been taken to have access to the necessary materials, in this case the most important are the video files of the security system of the establishment. The research begins with the review of the antecedents that have studied the proposed topic, finding a variety of them and in English and Spanish. Image processing was defined as the study variable, broken down into dimensions and indicators that help to operationalize it. In addition, the techniques and instruments have been taken into account in order to collect the necessary data for the ordering of the information, another important point in the development of the research are the ethical and administrative aspects.

Keywords: digital image, video image processing, video pixels, body gestures, convolutional neural networks.

I. INTRODUCCIÓN

Existen múltiples situaciones que diariamente se dan en la vida del ser humano que son repetitivas y que se pueden predecir con solo analizar los movimientos iniciales o anteriores a la acción a analizar, se sabe que hay múltiples formas de analizar estos movimientos, uno de ellos es el más usado de todos, se refiere a la observación con el ojo humano, pero también se sabe que esta forma de hacerlo no ofrece resultados 100% precisos, por eso en este proyecto de investigación se propone el desarrollo e implementación de un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes y video de los trabajadores de la empresa Servicios Generales y Transporte “Fresco Pez” en la ciudad de Chulucanas, con este software se espera como resultados que al procesar las imágenes de video de las acciones de los trabajadores muestre un panel de resultados describiendo las acciones que dichos trabajadores están realizando.

Una de las investigaciones que se toman como soporte para la realización de esta es Giant panda behaviour recognition using images (traducido al español, Reconocimiento del comportamiento del panda gigante mediante imágenes) de los autores Swarup, Chen, Hou, Que, Liu y Kong en el año 2021, que se traduce al español de esta forma, Reconocimiento del comportamiento del panda gigante mediante imágenes. En esta investigación se estudian los hábitos diarios del panda gigante, para asegurar la supervivencia del mismo ya que se encuentra en peligro de extinción. En el procesamiento de imágenes se pueden emplear varias técnicas, una de ellas es la sustracción de fondo, que en la investigación Procesamiento a imágenes de video para verificación de distanciamiento social durante la pandemia de la COVID-19 (Castro, Niño, Medina, Guevara & Gomez 2021) se afirma que es una de las técnicas más utilizadas en procesos de detección de personas, debido a la capacidad de separación de los objetos presentes en el cuadro principal de la imagen y el fondo de esta. Además, en esta investigación se usa el método de distancia euclidiana, en el

que se determina la separación entre puntos de un plano cartesiano basándose en el teorema de Pitágoras.

El desarrollo de esta investigación está compuesto por el estudio de varios indicadores o variables que son necesarios para obtener los resultados que se esperan, entre ellos se encuentran la detección de gestos corporales, reconocimiento facial, procesamiento de imágenes y comportamiento humano. El estudio se justificó metodológicamente ya que se utilizaron herramientas para la gestión de bibliografía, con el fin de almacenar y gestionar las investigaciones utilizadas como antecedentes. Además, se realizó un análisis y filtrado de cada uno de dichos antecedentes, seguido de una revisión de la metodología científica empleada en dichos estudios previos, esto con la ayuda de matrices de análisis, permitiendo conceptualizar la variable de estudio "Procesamiento de imágenes", con el fin de saber en qué nivel de investigación se encuentra la ciencia respecto al tema, como fueron evolucionando los trabajos de investigación en el tiempo y que dimensiones existen para tomarse como referencia de la investigación.

La investigación se justificó teóricamente ya que se realizó con el propósito de aportar al conocimiento que ya existe sobre el procesamiento de imágenes, además de que los resultados de la misma podrían incluirse en una propuesta para siguientes investigaciones. La justificación tecnológica se da ya que las tecnologías de información siguen en constante evolución y avance, además de que cada vez se exigen productos que sean modernos e innovadores de acuerdo a la época. Basándose en las ideas propuestas, se plantea las siguientes interrogantes: ¿Cómo un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes se relaciona con las acciones de los trabajadores de la empresa

Servicios Generales y Transporte "Fresco Pez" Chulucanas - 2021? Como preguntas específicas de acuerdo a las dimensiones de la variable se plantea ¿Cuál es la evaluación de los fondos de imagen en video por intensidad de color y posición?, ¿Cuál es la medición matricial por imagen de video por iluminación en la empresa Servicios Generales y Transporte Fresco Pez S.R.L.?, ¿Cómo medir la calidad de las imágenes de video de las cámaras de seguridad de la empresa Servicios Generales y Transporte Fresco Pez S.R.L.? y finalmente ¿Cómo medir la resolución de las imágenes de video

de las cámaras de seguridad de la empresa Servicios Generales y Transporte Fresco Pez S.R.L.?

De acuerdo a lo mencionado anteriormente se define como objetivo general de esta investigación implementar un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes y video de los trabajadores de la empresa Servicios Generales y Transporte “Fresco Pez” Chulucanas – 2021. Los objetivos específicos de la investigación de acuerdo a la planteado son evaluar los fondos de imagen en video por intensidad de color y posición de imagen, como segundo objetivo específico se define evaluar la medición matricial por el nivel de iluminación en la empresa Servicios Generales y Transporte Fresco Pez S.R.L., en tercer lugar, se tiene como objetivo específico medir la calidad de las imágenes de video de las cámaras de seguridad y finalmente, medir la resolución de las imágenes de video de las cámaras de seguridad de la empresa Servicios Generales y Transporte Fresco Pez S.R.L.

II. MARCO TEÓRICO.

En este capítulo se observan los estudios relacionados con el tema de investigación, tomando en cuenta libros, revistas, artículos que se desarrollaron en torno a temas como el procesamiento de imágenes, análisis de gestos corporales, análisis de conducta humana, entre otros. También se mostrará las teorías relacionadas al tema en cuestión, incluyendo teorías poco relacionadas con el rubro de los autores. Finalmente, se hará en desarrollo de conceptos importantes que se relacionan directamente con esta investigación, considerando la variedad de palabras que se deben explicar en este tema.

En los siguientes párrafos se realizará un resumen de los temas que engloban la investigación iniciando con el análisis del comportamiento humano, considerando en primer lugar a un elemento importante de los seres humanos, se refiere al lenguaje, que ha sido estudiado por distintos puntos de vista y disciplinas. En particular, el lenguaje escrito hoy es abordado desde la lingüística, la neurociencia, las ciencias de la visión, la psicología, y más recientemente la computación (Kamienkowski, Bengolea, Bianchi y Ferrer 2016).

Respecto al análisis del comportamiento, se puede afirmar que existen infinidad de gestos y actos básicos que juntándose empiezan a ganar complejidad y que son muy comunes en la vida diaria del ser humano. Estudios anteriores indican que se puede capturar el movimiento en tres métodos. El primer método está basado en aprendizaje, en este método se limita el número de movimientos utilizables. El segundo es un método sin modelo, que no utilizan información “a priori” para modelar el cuerpo humano y el tercero está basado en modelo, que hacen corresponder un modelo humano creado “a priori” con la información observada, en cualquier tipo de movimiento (Marcos, Pizarro, Marron y Mazo 2013).

Pranjal (Singapur, 2021) en su investigación Reconocimiento del comportamiento del panda gigante mediante imágenes tiene como objetivo utilizar un método completamente automático para clasificar el comportamiento del panda en cinco clases principales: caminar, sentarse, descansar, trepar y comer, así como determinar si un panda abre o cierra los ojos y la boca. Para examinar la solidez del método

propuesto, se estableció un gran conjunto de datos que consta de 10.804 imágenes de pandas de más de 218 pandas en cautiverio. Como conclusión de esta investigación, el autor afirma que monitorear el actuar de los pandas es de mucha importancia para la salud y conservación de la especie. Se ha desarrollado un método de reconocimiento del comportamiento de panda totalmente automático basado en el aprendizaje profundo. De esta forma, se pueden clasificar los cinco comportamientos principales de sentarse, caminar, descansar, trepar y comer y los dos sutiles movimientos faciales que abren y cierran los ojos y la boca.

En otras investigaciones se utilizan las redes neuronales convolucionales, que para los resultados se realizan una serie de experimentos para probar el método propuesto. Garcia (Lima, 2020) investigación titulada Identificación automática de las fases del gesto de recepción en el vóley mediante análisis de videos usando redes neuronales convolucionales propone un modelo de algoritmo que permite la identificación de manera automática del período de tiempo que ocurre durante la ejecución de la técnica de recepción de voleibol a partir del video. Se definen las fases a analizar, los módulos en los que se divide el modelo algorítmico, en este caso es el módulo de detección de jugador y módulo de clasificación de fases. En estos módulos se definen las arquitecturas de los modelos a utilizar. Finalmente se muestran los resultados obtenidos tras la implementación de los módulos mencionados. Los resultados muestran que el modelo implementado puede determinar con precisión los pasos de tiempo a nivel de cuadro con una precisión del 92,19%. Además, si ocurre un error de reconocimiento, la fase identificada por el modelo es adyacente a la fase real, lo que indica que el modelo puede capturar la naturaleza temporal de la fase.

Fernández (Arequipa, 2016) presenta un modelo para optimizar la detección de rostros en secuencias de video, basándose en un método para la detección de rostros en secuencias de video, a raíz de realizar mejoras en la iluminación, aumento de resolución y acelerar los cálculos computacionalmente con la unidad de procesamiento gráfico GPU. Como resultado se obtiene que, a pesar del aumento de resolución, todavía no es suficiente para este tipo de video y posteriores pruebas con varios videos

han identificado caras porque la resolución es muy baja porque es un video, podemos concluir que es difícil de hacer. Por distancia o nitidez reducida. Esto muestra que la implementación de la súper resolución mejora la calidad del video. Además, resulta que el mayor problema es la presencia de ruido de variación espacial entre cuadros debido a la presencia de animaciones de transición en la secuencia de video.

Bengoa (España, 2001) en su investigación reconocimiento automático de caras basado en el análisis de características físicas y la expresión, tiene como objetivo desarrollar un sistema y una metodología para el reconocimiento facial en imágenes digitales, teniendo en cuenta las distorsiones provocadas por las diferentes expresiones que pueden producir los rostros. Para ejecutar las pruebas, se creó una base de datos que tiene imágenes de caras de 20 personas (240 imágenes) que representan siete expresiones básicas: feliz, asustado, enojado, disgustado, neutral, triste, molesto y sorprendido. Realizar estandarización y organización de la matriz.

Para el desarrollo de esta investigación se deben tener en cuenta los gestos del ser humano y de las partes del cuerpo, Flores (Cuzco 2017) en su estudio titulado clasificación y reconocimiento de gestos estáticos realizados por la mano basado en el alfabeto dactilológico de la lengua de señas del Perú aplicando redes profundas bajo características invariantes, presenta una clasificación de gestos de la mano humana basado en el alfabeto dactilológico de la lengua de señas, afirma que el reconocimiento es un ámbito muy activo de investigación. Tiene como objetivo desarrollar una arquitectura de una red neuronal convolucional para clasificar un conjunto de 24 gestos obtenidos de la base de datos que se desarrolló en la investigación. Concluyó que su proyecto de investigación no solo apoya a muchos investigadores en la línea de investigación, sino que también ayuda a desarrollar futuros proyectos relacionados con el procesamiento de imágenes digitales, el reconocimiento de gestos con las manos y el aprendizaje profundo.

El autor Sánchez (2015) en su investigación reconocimiento automático de actividades humanas basado en el análisis de secuencias de imágenes, afirma que el seguimiento de personas en secuencia de video y la interpretación automática de las actividades que desarrollan es una tarea que tiene una gran variedad de aplicaciones en áreas

como vídeo-vigilancia, interacción persona-ordenador, análisis de la técnica deportiva y otros más. Para realizar esto se propone una combinación de métodos avanzados de seguimiento visual, reconocimiento de patrones y aprendizaje permitirá desarrollar métodos automáticos eficaces y eficientes de reconocimiento de actividades humanas. Como resultado, se puede verificar que el sistema propuesto proporciona resultados similares en términos de exactitud y precisión para el seguimiento entre 2 y 15 sujetos. En trabajos futuros, nuestro objetivo es comparar nuestro sistema con otros sistemas propuestos en la literatura actual sobre secuencias de imágenes.

Los autores Cadena, Montaluisa y otros (2017) en su estudio el reconocimiento facial basado en imágenes, afirman que lo mencionado es de suma importancia en diferentes ámbitos, mayormente en seguridad, para dar acceso a sitios privados o registro de llegada y salida. La extracción de la mitad del rostro de cada persona proporciona información para la detección de rostros. El intercambio también involucra los diferentes procesos, etapas y métodos de extraer las características del funcionamiento del sistema de reconocimiento facial. Incluyendo los aspectos positivos y negativos de su implementación, pros y contras, y los estándares de las personas al respecto. Como resultado, se puede afirmar que el reconocimiento facial basado en imágenes es muy importante. Según los datos de la encuesta realizada a 50 personas de la ciudad de Latacunga, el 80% concuerda que es de suma importancia, además el 90% afirmó que es de mucha ayuda en los casos policiales en el país. En conclusión, este estudio encontró que ninguna tecnología de reconocimiento facial cumple con todas las expectativas necesarias.

En la investigación de Yamashita, Hirakawa y Fujiyoshi (2019) titulada reconocimiento de imágenes basado en aprendizaje profundo para conducción autónoma se observa que varias tareas de reconocimiento de imágenes se manejaron en el campo del reconocimiento de imágenes antes de 2010 mediante la combinación de características locales de imágenes diseñadas manualmente por los investigadores y el método de aprendizaje automático. Sin embargo, después de ingresar al 2010, se han propuesto muchos métodos de reconocimiento de imágenes que utilizan el

aprendizaje profundo. Este artículo explica cómo se aplica el aprendizaje profundo al campo del reconocimiento de imágenes y también explicará las últimas tendencias de la conducción autónoma basada en el aprendizaje profundo.

También, se han tomado en cuenta investigaciones sobre el comportamiento humano ya sea laboral, en redes sociales u otros ámbitos. Entre esas investigaciones se encuentra el estudio titulado: “Gestión de recursos humanos y comportamiento laboral de los colaboradores del Hotel Jazmín, Lince-2019”, tuvo como objetivo determinar la relación que existe entre la Gestión de recursos humanos y comportamiento laboral de los colaboradores del Hotel Jazmín, Lince 2019. Para obtener resultados se usó la técnica de la encuesta para medir el comportamiento laboral de los involucrados. Estos resultados se validaron mediante el juicio de expertos, para dar confiabilidad se realizó una prueba piloto, cuyos resultados se sometieron al Alfa de Cronbach. Luego de la investigación realizada se concluyó que la relación entre la gestión de recursos humanos y el comportamiento laboral es alta, ya que una adecuada gestión de los recursos va a modificar el comportamiento del personal de la empresa de buena manera, lo que le permitirá brindar una adecuada calidad de servicio para con los clientes (Aybar 2019).

Existen conceptos dentro de la investigación que se van a definir a continuación, entre ellos están lenguaje corporal, procesamiento digital de imágenes, hardware y software. Lenguaje corporal abarca un conjunto de conceptos que ayudarán al desarrollo de esta investigación, uno de ellos es el lenguaje corporal, que estudia el significado comunicativo de los gestos aprendidos no orales, otro concepto que se incluye en la investigación es la forma de comunicación no verbal ejecutada con alguna parte del cuerpo y producida por el movimiento de las articulaciones, músculos de brazos, manos o cabeza a la que se conoce como gestos (Moreno 2015).

La imagen digital está compuesta por una cantidad definida de píxeles, estos contienen un valor y una posición particular. La definición de video también incluye a los píxeles, y estos están formados por un número de bits (Medina, 2018).

El procesamiento digital de imágenes es el conjunto de técnicas que se aplican a las imágenes digitales con el objetivo de mejorar su calidad, también existe el análisis de

imágenes incluye aquellas técnicas cuyo principal objetivo es facilitar la búsqueda e interpretación de información contenida en ellas. El procesamiento de imágenes tiene definido su origen en una imagen y su fin llega junto con la imagen procesada, el valor de la calidad de la imagen procesada se encuentra en relación a la imagen de entrada. En este tema se observan varias técnicas, tales como la mejora de la calidad, técnicas de restauración, compresión y segmentación (Mendoza 2009).

Medina (Trujillo, 2018) define la resolución de imagen como la facultad para distinguir dos puntos muy cercanos antes de que se fusionen en uno, porque una imagen digital consta de puntos discretos que se consideran un continuo a una cierta distancia.

III. METODOLOGÍA

3.1. Tipo y diseño de investigación

Para el desarrollo de este estudio se ha definido que el tipo de investigación sea descriptivo. La investigación descriptiva se utiliza para analizar cómo se ve un fenómeno, sus componentes y cómo se manifiesta. Permiten concretar el fenómeno en estudio midiendo una o más de sus propiedades (Vásquez).

El diseño de investigación definido para el desarrollo del tema propuesto es un diseño **no experimental transeccional**, se llama estudio no experimental porque el investigador no controla las variables de investigación, también se denomina estudio transeccional, ya que recolecta datos de un momento y una vez. Su propósito es describir variables y analizar su incidencia e interrelaciones en un momento determinado (Hernández). Siendo x_1 , x_2 , x_3 y x_4 dimensiones de la variable procesamiento de imágenes.

Diseño:

G: Imágenes de video captadas por el sistema de seguridad de la empresa relacionadas con los gestos corporales y acciones repetitivas de los trabajadores.

G: X1 ----- O1

X1: Fondo de imagen.

O1: Se observa la posición de imagen y la intensidad del color utilizando una ficha de registro.

G: X2 ----- O2

X2: Medición matricial.

O2: Se observa el nivel de iluminación utilizando una ficha de registro.

G: X3 ----- O3

X3: Calidad de imagen.

O3: Se observa la cantidad de pixeles por pulgada utilizando una ficha de registro. G: X4 ----- O4

X4: Resolución de imagen

O4: Se observa el tamaño y nivel de pixeles utilizando una ficha de registro.

3.2. Variables y operacionalización

Tabla 1. Matriz de operacionalización de la variable.

Variable	Definición conceptual	Definición operacional	Dimensiones	Indicadores	Escala de medición
Procesamiento de imágenes	Se refiere a procesar las imágenes de manera digital mediante un computador. Es un tema muy amplio, en el que se incluyen estudios de física, matemáticas, ingeniería eléctrica, computación (Medina 2018).	Por cada dimensión de la variable se realizará una ficha de registro, además del uso de videos de seguridad, las imágenes serán analizadas en fases.	Fondo de imagen	Posición de imagen	Escala/ Discreta
				Intensidad de color	
			Medición matricial	Nivel de iluminación	
			Calidad de imagen	Cantidad de pixeles por pulgada	
			Resolución de imagen	Tamaño de imagen	
	Nivel de pixeles				

Elaborado por los autores.

3.3. Población, muestra y muestreo

Según López (Bolivia, 2004), población es un grupo de individuos o elementos de los que quiere aprender sobre algo en el estudio y la muestra es una subpoblación o una parte del mundo en que se llevará a cabo el estudio.

Población: En el estudio a realizar se tomará como población a las imágenes y videos de 3 cámaras de vigilancia y de 10 operarios como trabajadores con acciones y trabajo repetitivo.

- **Criterios de inclusión:** Para la inclusión se tomarán en cuenta personas de mayor edad que laboran y aceptan que se utilice desde las cámaras de la empresa, informándoles que se respetará su anonimato.
- **Criterios de exclusión:** Personas menores de 18 años o que no acepten que se utilice los videos donde aparecen ellos serán excluidos de la población.

Muestra: Será una muestra por conveniencia de imágenes y video entre los meses de julio, agosto y septiembre de 2021, en un lapso de tiempo correspondiente a ocho horas diarias (6:30 a.m. – 2:30 p.m.), durante toda la semana. La muestra de esta investigación estará compuesta específicamente por las imágenes y videos de los trabajadores y propietario de la empresa Servicios Generales y Transporte Fresco Pez S.R.L. que tiene su establecimiento comercial ubicado en el Jr. Colón 699 del Distrito de Chulucanas.

3.4. Técnicas e instrumentos de recolección de datos

La técnica definida para la recopilación de datos necesarios de este estudio es la observación directa. La observación es el proceso de mirar detenida o profundamente, en sentido amplio, el experimento, el proceso de someter conductas de algunas cosas o condiciones manipuladas de acuerdo a ciertos principios para llevar a cabo un cabo la observación. Observación significa también el conjunto de cosas observadas, el conjunto de datos y conjunto de fenómenos (Pardinas). La observación directa es cuando el investigador entra en contacto personalmente con el hecho o situación que está tratando de evaluar (Diaz, 2011).

Los instrumentos que se van a utilizar serán las fichas de registro por cada dimensión de las variables de esta investigación, en esas fichas se registrará los resultados de las pruebas realizadas. Las fichas de registro se llenarán con los datos que los investigadores obtengan usando la técnica ya mencionada, se asignará un identificador a los archivos de video formado por la fecha y hora del video (ej. 1107210929). Se describirán los momentos precisos de los videos que se utilizarán en el desarrollo del estudio. Finalmente, según la formulación de cada dimensión se registran los resultados obtenidos.

3.5. Procedimientos

Como primera fase se tiene la captura o adquisición de las imágenes, que se obtienen de las cámaras instaladas en el establecimiento, la segunda fase se refiere al preprocesamiento, cuyo proceso es el de reparar la imagen, en tercer lugar se encuentra la segmentación, que divide la imagen en objetos de suma importancia para la investigación, luego de ello, seguiría la descripción, que se refiere a obtener características que ayuden a definir objetos u otros, y finalmente la comparación, que clasifica los objetos según las características de los mismos.

Con autorización escrita y firmada por el Gerente y propietario de la empresa Servicios Generales y Transporte Fresco Pez S.R.L. se realizará el traslado de los archivos de video del grabador del sistema de seguridad del establecimiento a una unidad de almacenamiento externo, de estos archivos se hará una selección de los mismos tomando en cuenta el contenido de los mismos, dicho contenido deberá ser netamente para fines de estudio, los rostros de las personas se mostrarán de forma distorsionada en caso se llegara a publicar la investigación. El consentimiento informado con la firma del gerente y propietario de la empresa Servicios Generales y Transporte Fresco Pez S.R.L. se encuentra en los anexos de esta investigación.

3.6. Método de análisis de datos

Para esta investigación se usará un análisis descriptivo ya que es un método preciso de recopilar información y luego proceder a describir las relaciones que existen entre los datos.

Frecuencia absoluta (Matos, Contreras y Olaya, 2020), el valor de esta es el número de veces que los datos de la muestra son de una determinada clase. Se representa por f_i . Para una muestra de tamaño "n" y con "k" clases se cumple que la sumatoria de todas las frecuencias absolutas es igual a "n":

$$\sum_{i=1}^k f_i = n$$

Siendo:

F_i = Valor de la frecuencia absoluta en la i-esima clase.

n = Tamaño de la muestra.

k = Números de clases.

La frecuencia absoluta se utilizará en relación a la cantidad de veces que una acción (propia del puesto de trabajo) se repite en un determinado tiempo.

Frecuencia relativa (Matos, Contreras y Olaya, 2020), su valor es el porcentaje de la muestra está contenida en una determinada clase. Se representa por h_i y se obtiene al dividir la frecuencia absoluta por el tamaño de la muestra:

$$h_i = \frac{f_i}{n} * 100\% \text{ (En porcentaje)}$$

En el caso de la frecuencia relativa, se usará para indicar el porcentaje de la muestra en la que se realizan las acciones repetitivas evaluadas en la frecuencia absoluta.

3.7. Aspectos éticos

Para el avance de esta investigación se deben tener en cuenta los siguientes aspectos éticos:

- El respeto por la información recolectada, partiendo desde las referencias propuestas y siguiendo las normas de documentación.
- La información que se obtenga será utilizada sólo con fines de investigación.
- Respeto al código de ética del colegio de ingenieros del Perú y de la investigación en la universidad Cesar Vallejo.
- La información recopilada de diferentes autores será debidamente citada, siguiendo el estándar ISO 690.

IV. RESULTADOS

Los resultados de obtuvieron usando la técnica de observación directa utilizando como instrumento fichas de registro para cada dimensión de la variable de estudio.

Figura 1. Ficha de registro N°1.

Variable		Indicador	Medida		Fórmula	
Procesamiento de imagen		Posición de imagen Intensidad de color	Deciles y Cuartiles Escala RGB			
#	Fecha de revisión	Identificación del procesamiento	Tema (Imágenes del sistema de seguridad)	Contenido específico	Coordenadas	Intensidad de color
1	13-10-2021	foto129	Venta de producto	Pesando	X: 312; Y: 300	211,211,211
2	13-10-2021	foto131	Venta de producto	Pesando	X: 375; Y: 402	143,143,143
3	13-10-2021	foto137	Venta de producto	Pesando	X: 338; Y: 236	186,186,186
4	13-10-2021	foto333	Venta de producto	Cobrando	X: 332; Y: 435	62,62,62
5	13-10-2021	foto368	Venta de producto	Cobrando	X: 278; Y: 398	134,134,134
6	13-10-2021	foto1583	Venta de producto	Cobrando	X: 290; Y: 517	56,56,56
7	13-10-2021	foto1767	Venta de producto	Cobrando	X: 300; Y: 406	76,76,76
8	13-10-2021	imgs2765	Venta de producto	Cobrando	X: 258; Y: 482	73,73,73
9	13-10-2021	photo3212	Venta de producto	Cobrando	X: 216; Y: 472	41,41,41
10	13-10-2021	image324	Venta de producto	Pesando	X: 304; Y: 214	46,46,46
11	13-10-2021	foto6543	Venta de producto	Pesando	X: 305; Y: 235	248,248,248
12	13-10-2021	foto4653	Venta de producto	Pesando	X: 306; Y: 241	237,237,237
13	13-10-2021	imgs0	Venta de producto	Existencia de clientes	X: 147; Y: 212	62,62,62
14	13-10-2021	imgs9	Venta de producto	Existencia de clientes	X: 147; Y: 212	62,62,62
15	13-10-2021	imgs15	Venta de producto	Existencia de clientes	X: 147; Y: 212	62,62,62
16	13-10-2021	imgs21	Venta de producto	Existencia de clientes	X: 147; Y: 212	62,62,62
17	13-10-2021	imgs1118	Venta de producto	Ausencia de clientes	X: 147; Y: 212	87,87,87
18	13-10-2021	imgs1131	Venta de producto	Ausencia de clientes	X: 147; Y: 212	87,87,87

Se han realizado la creación de dos modelos utilizando redes neuronales y que han sido entrenados con un total de 5190 imágenes, 2590 imágenes para el modelo principal que logra realizar el reconocimiento entre dos gestos observados; el modelo secundario está dedicado a confirmar la existencia de clientes en el establecimiento, este modelo se entrenó con 2600 imágenes. En la ficha de registro se han tomado los datos de 20 imágenes que se han usado para comprobar la eficacia de los modelos. Se puede afirmar que debido a que la posición del equipo (balanza, caja de dinero) varía, las coordenadas de las acciones **pesando** y **cobrando** no son las mismas en todas las imágenes, la

posición de los equipos no es constante ya que estos son colocados por los trabajadores de la empresa, además, el tamaño de imagen no es el mismo en todas. Por otro lado, las imágenes utilizadas para la comprobación del modelo secundario si tienen las mismas coordenadas ya que los clientes siempre están ubicados en la misma ubicación.

Figura 2. Ficha de registro N°2.

Variable		Indicador	Medida		Fórmula
Procesamiento de imagen		Nivel de iluminación	Lux (Lx)		Lumen / m ²
#	Fecha de revisión	Identificación del procesamiento	Tema (Imágenes del sistema de seguridad)	Contenido específico	Lux(lx)
1	13-10-2021	foto129	Venta de producto	Pesando	75 lx
2	13-10-2021	foto131	Venta de producto	Pesando	75 lx
3	13-10-2021	foto137	Venta de producto	Pesando	75 lx
4	13-10-2021	foto333	Venta de producto	Cobrando	75 lx
5	13-10-2021	foto368	Venta de producto	Cobrando	75 lx
6	13-10-2021	foto1583	Venta de producto	Cobrando	75 lx
7	13-10-2021	foto1767	Venta de producto	Cobrando	75 lx
8	13-10-2021	imgs2765	Venta de producto	Cobrando	75 lx
9	13-10-2021	photo3212	Venta de producto	Cobrando	75 lx
10	13-10-2021	image324	Venta de producto	Pesando	75 lx
11	13-10-2021	foto6543	Venta de producto	Pesando	75 lx
12	13-10-2021	foto4653	Venta de producto	Pesando	75 lx
13	13-10-2021	imgs0	Venta de producto	Existencia de clientes	75 lx
14	13-10-2021	lms9	Venta de producto	Existencia de clientes	75 lx
15	13-10-2021	lms15	Venta de producto	Existencia de clientes	75 lx
16	13-10-2021	lms21	Venta de producto	Existencia de clientes	75 lx

En la segunda ficha de registro se puede observar una constante con respecto al nivel de iluminación, ya que este se mide de acuerdo a la división del lumen de cada luminaria (foco) y el área del establecimiento. Para este caso se cuenta con 3 luminarias de 4000 lumen cada una y un espacio de 160 m², dando como resultado un total de 75 lx (lux).

Figura 3. Ficha de registro N°3.

Variable		Indicador	Medida		Fórmula
Procesamiento de imagen		Cantidad de pixeles por pulgada	Píxeles (PX)		
#	Fecha de revisión	Identificación del procesamiento	Tema (Imágenes del sistema de seguridad)	Contenido específico	Número de pixeles
1	13-10-2021	foto129	Venta de producto	Pesando	128,44 ppp
2	13-10-2021	foto131	Venta de producto	Pesando	50,04 ppp
3	13-10-2021	foto137	Venta de producto	Pesando	128,44 ppp
4	13-10-2021	foto333	Venta de producto	Cobrando	128,44 ppp
5	13-10-2021	foto368	Venta de producto	Cobrando	105,32 ppp
6	13-10-2021	foto1583	Venta de producto	Cobrando	128,44 ppp
7	13-10-2021	foto1767	Venta de producto	Cobrando	128,44 ppp
8	13-10-2021	imgs2765	Venta de producto	Cobrando	128,44 ppp
9	13-10-2021	photo3212	Venta de producto	Cobrando	128,44 ppp
10	13-10-2021	image324	Venta de producto	Pesando	80,79 ppp
11	13-10-2021	foto6543	Venta de producto	Pesando	80,68 ppp
12	13-10-2021	foto4653	Venta de producto	Pesando	81,79 ppp
13	13-10-2021	imgs0	Venta de producto	Existencia de clientes	128,44 ppp
14	13-10-2021	imgs9	Venta de producto	Existencia de clientes	128,44 ppp
15	13-10-2021	imgs15	Venta de producto	Existencia de clientes	128,44 ppp
16	13-10-2021	imgs21	Venta de producto	Existencia de clientes	128,44 ppp

La ficha de registro número 3 toma en cuenta la cantidad de pixeles por pulgada, este dato se obtiene en relación a la resolución o tamaño de la pantalla del equipo utilizado para la visualización (Laptop) con la altura y anchura de la imagen. Se observa que en las imágenes utilizadas para la comprobación del segundo modelo la cantidad pixeles por pulgada es la misma ya que dichas imágenes se utilizaron con el tamaño original, por otro lado, sucede lo contrario con las imágenes relacionadas al modelo principal (desde la imagen 1 hasta la 12) ya que estas imágenes fueron recortadas por los autores para una mejor visualización de los gestos realizados en ellas. Dicho recorte es de suma importancia para el entrenamiento del modelo principal debido a que aporta una mayor precisión al momento de probar el modelo.

Figura 4. Ficha de registro N°4.

Variable		Indicador	Medida		Fórmula	
Procesamiento de imagen		Tamaño de imagen Nivel de pixeles	Píxeles (PX)		Tamaño de px -> $(\text{ppp} * 2.54)/100$	
#	Fecha de revisión	Identificación del procesamiento	Tema (Imágenes del sistema de seguridad)	Contenido específico	Cantidad de píxeles	Tamaño de píxeles
1	13-10-2021	foto129	Venta de producto	Pesando	1843200 px	3.26 mm
2	13-10-2021	foto131	Venta de producto	Pesando	277992 px	12.71 mm
3	13-10-2021	foto137	Venta de producto	Pesando	1843200 px	3.26 mm
4	13-10-2021	foto333	Venta de producto	Cobrando	1843200 px	3.26 mm
5	13-10-2021	foto368	Venta de producto	Cobrando	1185280 px	2.68 mm
6	13-10-2021	foto1583	Venta de producto	Cobrando	1843200 px	3.26 mm
7	13-10-2021	foto1767	Venta de producto	Cobrando	1843200 px	3.26 mm
8	13-10-2021	imgs2765	Venta de producto	Cobrando	1843200 px	3.26 mm
9	13-10-2021	photo3212	Venta de producto	Cobrando	1843200 px	3.26 mm
10	13-10-2021	image324	Venta de producto	Pesando	726138 px	2.05 mm
11	13-10-2021	foto6543	Venta de producto	Pesando	726291 px	2.04 mm
12	13-10-2021	foto4653	Venta de producto	Pesando	747355 px	2.08 mm
13	13-10-2021	imgs0	Venta de producto	Existencia de clientes	1843200 px	3.26 mm
14	13-10-2021	imgs9	Venta de producto	Existencia de clientes	1843200 px	3.26 mm
15	13-10-2021	imgs15	Venta de producto	Existencia de clientes	1843200 px	3.26 mm
16	13-10-2021	imgs21	Venta de producto	Existencia de clientes	1843200 px	3.26 mm
17	13-10-2021	imgs1118	Venta de producto	Ausencia de clientes	1843200 px	3.26 mm
18	13-10-2021	imgs1131	Venta de producto	Ausencia de clientes	1843200 px	3.26 mm

En la ficha de registro 4 se observan los datos respecto al tamaño de la imagen y de los píxeles de cada una de las imágenes en cuestión. En este caso sucede lo mismo que la ficha anterior ya que depende principalmente del tamaño de las imágenes, se puede observar que el tamaño original de las imágenes es de 1843200 píxeles y que el tamaño de los píxeles en cada una de estas imágenes es de 3.26 mm, en el resto de imágenes estos valores varían ya que el recorte de estas no se realizó de manera uniforme.

Tabla 2. Comprobación del modelo principal.

N°	Nombre	Resultado esperado	Resultado obtenido	
1	foto129	Pesando	Cobrando	FALSO
2	foto131	Pesando	Pesando	VERDADERO
3	foto137	Pesando	Pesando	VERDADERO
4	foto333	Cobrando	Cobrando	VERDADERO
5	foto368	Cobrando	Cobrando	VERDADERO
6	foto1583	Cobrando	Cobrando	VERDADERO
7	foto1767	Cobrando	Cobrando	VERDADERO
8	foto4653	Pesando	Pesando	VERDADERO
9	foto6543	Pesando	Cobrando	FALSO
10	im40	Cobrando	Cobrando	VERDADERO
11	im41	Cobrando	Cobrando	VERDADERO
12	im42	Cobrando	Cobrando	VERDADERO
13	im43	Cobrando	Cobrando	VERDADERO
14	im44	Cobrando	Cobrando	VERDADERO
15	im45	Cobrando	Cobrando	VERDADERO
16	im46	Cobrando	Cobrando	VERDADERO
17	im47	Cobrando	Cobrando	VERDADERO
18	im48	Cobrando	Cobrando	VERDADERO
19	im49	Cobrando	Cobrando	VERDADERO
20	Im50	Cobrando	Cobrando	VERDADERO
21	Im51	Cobrando	Cobrando	VERDADERO
22	im668	Cobrando	Cobrando	VERDADERO
23	im669	Cobrando	Cobrando	VERDADERO
24	im670	Cobrando	Cobrando	VERDADERO
25	im671	Cobrando	Cobrando	VERDADERO
26	im672	Cobrando	Cobrando	VERDADERO
27	im673	Cobrando	Cobrando	VERDADERO
28	im674	Cobrando	Cobrando	VERDADERO
29	im675	Cobrando	Cobrando	VERDADERO
30	image324	Pesando	Pesando	VERDADERO
31	image1485	Pesando	Pesando	VERDADERO
32	imgs2765	Cobrando	Cobrando	VERDADERO
33	imgs4345	Pesando	Pesando	VERDADERO
34	imgs4346	Pesando	Pesando	VERDADERO
35	imgs4347	Pesando	Pesando	VERDADERO
36	imgs4348	Pesando	Pesando	VERDADERO
37	imgs4349	Pesando	Pesando	VERDADERO
38	imgs4350	Pesando	Pesando	VERDADERO
39	imgs4351	Pesando	Pesando	VERDADERO
40	imgs4352	Pesando	Pesando	VERDADERO
41	imgs4353	Pesando	Pesando	VERDADERO
42	imgs4354	Pesando	Pesando	VERDADERO
43	imgs4355	Pesando	Pesando	VERDADERO
44	imgs4356	Pesando	Pesando	VERDADERO
45	imgs4357	Pesando	Pesando	VERDADERO
46	imgs4358	Pesando	Pesando	VERDADERO
47	imgs4359	Pesando	Pesando	VERDADERO
48	imgs4360	Pesando	Pesando	VERDADERO
49	imgs4361	Pesando	Pesando	VERDADERO
50	photo3212	Cobrando	Cobrando	VERDADERO

Elaborado por los autores.

Se puede observar que en 48 de las 50 pruebas se obtiene un resultado satisfactorio, llegando así a calcular un 96% de eficacia en el reconocimiento de gestos corporales por parte del modelo principal.



Figura 5. Imágenes con resultado erróneo.

Al realizar un análisis detallado de la imagen se llega a un punto en donde debido a la cercanía de la mano de uno de los trabajadores a la caja contenedora de dinero genera una confusión para el modelo, y por ende un resultado erróneo al momento de la ejecución del mismo. En otra imagen se puede observar que hay un claro aumento de iluminación lo que genera distorsión al momento de realizar el procesamiento.

Tabla 3. Comprobación del modelo secundario.

Prueba	Archivo	Resultado esperado	Resultado obtenido
1	imgs0	Si hay clientes	Si hay clientes
2	limgs9	Si hay clientes	Si hay clientes
3	limgs15	Si hay clientes	Si hay clientes
4	limgs21	Si hay clientes	Si hay clientes
5	imgs1118	No hay clientes	No hay clientes
6	imgs1131	No hay clientes	No hay clientes
7	imgs1133	No hay clientes	No hay clientes
8	imgs1135	No hay clientes	No hay clientes

Elaborado por los autores.

En la tabla 6 se observa que los resultados obtenidos en las pruebas del modelo secundario son en su totalidad correctos, comprobando así que el modelo tiene una precisión del 100%.

Es necesario hacer mención al proceso de entrenamiento del modelo principal se realizó en un intervalo de tiempo de 18 minutos y 30 segundos, tomando en cuenta que los datos de entrenamiento eran un total de 2590. Para el segundo modelo se tardó 31 minutos y 42 segundos, el número de imágenes es similar al modelo principal (2600), tomando en cuenta que todas las imágenes de este modelo se usaron con el tamaño original a diferencia del modelo principal donde se recortaron todas las imágenes. Cabe resaltar que este tiempo puede ser mayor o menor según los recursos informáticos con los que se cuenta. En este caso se utilizó una laptop de marca **Acer** con el sistema operativo Windows 11, 12Gb de memoria RAM, 6Gb de tarjeta gráfica, procesador Intel(R) Core (TM) i5 de 11th generación y una unidad de almacenamiento solida de 236.8 Gb.

Tabla 4. Tabla de frecuencias.

Dato	F	Fr	%
Pesando	22	0.379310345	37.9310345
Cobrando	26	0.448275862	44.8275862
Error	2	0.034482759	3.44827586
No Clientes	4	0.068965517	6.89655172
Clientes	4	0.068965517	6.89655172
Total	58	1	100

Elaborado por los autores.

En esta tabla se han definido en la columna de datos, los ítems referentes a los resultados que se muestran en cada modelo creado, los datos de **Pesando**, **Cobrando** y **Error** corresponden al modelo principal, en donde de 50 pruebas, se encontró error en 2 de ellas. Todo lo contrario, sucede con el modelo secundario, en el cual hay una efectividad del 100%. En esta tabla se han encontrado la frecuencia relativa y absoluta de un total de 58 datos.

V. DISCUSIÓN

El objetivo general de esta investigación es implementar un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes y video de los trabajadores de la empresa Servicios Generales y Transporte “Fresco Pez” Chulucanas – 2021, para ello se revisaron estudios que aportan información importante para el desarrollo de dicho sistema como Pranjali (Singapur, 2021) que utilizó un total de 10804 imágenes para lograr la detección de 9 gestos corporales por parte de 218 pandas gigantes, otro de los autores (Bengoa, España, 2001) utilizó un total de 240 imágenes de 20 rostros de personas para lograr la identificación de 7 expresiones faciales. Por otro lado, este tema para los autores Cadena, Montaluisa y otros (2017) es fundamental en el ámbito de seguridad, dándole mayor énfasis al control de entrada y salida de lugares privados con reconocimiento facial.

Con respecto a las investigaciones mencionadas anteriormente, para esta investigación se propuso la creación de dos modelos desarrollados en el editor de texto **Sublime Text** usando el lenguaje de programación **Python**, para el proceso de entrenamiento del modelo principal (que diferencia entre las acciones **pesando** y **cobrando**) se usaron 2590 imágenes, dichas imágenes se tomaron de los videos de las cámaras de seguridad del establecimiento, los videos fueron procesados y se obtuvo cada fotograma de los videos seleccionados, cada imagen que se tomó tiene el formato **png**, en blanco y negro, y recortadas enfocándose en la acción a evaluar. El modelo principal obtuvo un 96% de precisión, para la creación del modelo secundario (que define si en una imagen existen o no clientes) se utilizaron 10 imágenes más que en el modelo principal (2600), estas imágenes se utilizaron en el tamaño original obteniendo una precisión del 100%. La precisión de ambos modelos puede variar de acuerdo a la cantidad de pruebas, hasta el momento la precisión del modelo secundario no ha variado, pero puede haber alguna diferencia considerando lo sucedido con el modelo principal.

Los autores Marcos, Pizarro, Marron y Mazo (2013) proponen el desarrollo de un método sin modelo, que solo toma en cuenta los movimientos captados con la técnica de observación humana, en el caso de esta investigación se ha definido la creación de dos modelos, pero existe una similitud entre dicho antecedente y la investigación en curso, ya que para la creación de los datos de entrenamiento se basaron en la observación con el ojo humano, y también el uso de herramientas para editar los archivos necesarios para el entrenamiento de los modelos creados. El autor Fernández (Arequipa, 2016) realizó un modelo para optimizar la detección de rostros en video, con base en un método que se centra en realizar mejoras en la iluminación y aumento en la resolución. En esta investigación no se han considerado las mejoras en la iluminación ni en la resolución, sin embargo, se plantea realizar un comentario en el apartado de recomendaciones para que en investigaciones futuras tomen en cuenta ambos temas para obtener mejores resultados.

En relación al primer objetivo específico que fue evaluar los fondos de imagen en video por intensidad de color y posición de imagen que de acuerdo con Oviedo (2004) se denomina con el nombre de “fondo” al elemento de homogeneidad que ofrece un grado de información constante e invariable que le permite al sujeto tener una impresión sensorial fácilmente constatable y que se mide respecto a coordenadas e intensidad de color, se obtuvo que la posición de la imagen no siempre va a ser la misma ya que depende de que el equipo en cuestión siempre esté ubicado en el mismo lugar, acción que en este caso no sucede, además, las imágenes de entrenamiento no tienen el mismo tamaño o dimensiones ya que fueron recortadas por los autores. En el tema de intensidad de color se usó la escala RGB como medida, pero los tres valores que se obtienen son el mismo ya que las imágenes se encuentran en escala de grises.

Como segundo objetivo específico se plantea evaluar la medición matricial por el nivel de iluminación en la empresa Servicios Generales y Transporte Fresco

Pez S.R.L., que según Condes (2018) en la medición matricial el procesador de la cámara divide la escena en una serie de zonas que ocupan todo el cuadro y calcula la exposición correcta en base a todas ellas haciendo un promedio. Por esta razón, es un modo de medición que se utiliza en una gran cantidad de situaciones por defecto. En este caso se mide por el nivel de iluminación, que como se mencionó antes está definido de manera estática por 3 luminarias de 4000 lumen cada una y un área de 160 m² dando como resultado un total de 75 lx (lux). En este ámbito no se ha considerado la iluminación exterior proveniente del sol y que varía con el clima, para ello se plantea una recomendación.

El tercer objetivo específico es medir la calidad de las imágenes de video de las cámaras de seguridad, que de acuerdo al IAEA (2021) no existe una definición objetiva de la calidad de imagen, es más bien cuestión del juicio subjetivo del observador. En medicina nuclear, la base de la calidad de imagen es la capacidad del sistema de imagen para detectar diferencias de captación de un radiofármaco entre una zona de lesión y las zonas adyacentes, para esta dimensión se planteó una medición de acuerdo a la cantidad de píxeles por pulgada (ppp) obteniendo como resultado que este ítem depende de la altura y anchura de la imagen en cuestión, además de que se toma en cuenta la cantidad de pulgadas de la pantalla en la que se visualiza la imagen (en este caso 15 pulgadas).

Finalmente, el cuarto objetivo específico se refiere a medir la resolución de las imágenes de video de las cámaras de seguridad de la empresa Servicios Generales y Transporte Fresco Pez S.R.L. que según Medina (2018) es la capacidad de un detector para discriminar detalles o, más exactamente, la capacidad para distinguir dos puntos muy próximos antes de que se fundan en uno solo, dado que las imágenes digitales están compuestas de puntos discontinuos que se perciben como un continuo a una determinada distancia, se mide respecto a la cantidad y tamaño de píxeles de la imagen. En este caso

el tamaño de un pixel se relaciona con los datos del objetivo específico anterior (píxeles por pulgada).

En la tabla de comprobación del modelo principal (Tabla 5) se han realizado 50 pruebas con imágenes diferentes en tamaño y posición de objetos obteniendo un resultado bastante satisfactorio (96% de efectividad), este porcentaje de efectividad se tornaba menor en una primera instancia, ya que se realizó una prueba con 12 datos, en los que se obtuvo un 88% de efectividad, lo que lleva a la conclusión de que a más pruebas realizadas la efectividad del modelo principal aumentará. En el segundo modelo la efectividad no varía (100%), ya que el entrenamiento se basa en la existencia o ausencia de clientes en las imágenes. Los resultados obtenidos de las pruebas realizadas al modelo secundario (Tabla 6) se basan en 8 imágenes con un 50% de imágenes con ausencia de clientes y el otro 50% de imágenes con al menos un cliente enfocado.

VI. CONCLUSIONES

1. Luego de la obtención de los resultados de esta investigación, se realizó la comparación de estos con los de los antecedentes planteados, y se han encontrado similitudes y diferencias considerables. Se midió el fondo de imagen tomando en cuenta la posición de imagen que se evaluó de acuerdo a las coordenadas centrales de la acción que se observa en la imagen, además, en esas mismas coordenadas se midió la intensidad de color en escala RGB. Se llegó a la conclusión de que las acciones evaluadas se realizan mayormente en la misma posición, ya que los equipos de trabajo utilizados tienen aspectos que no varían, entre ellos la ubicación de los mismos.
2. Se tomó en cuenta la iluminación del establecimiento, cuyo resultado es el mismo para todas las pruebas realizadas, ya que no se abarcan los factores correspondientes a la iluminación externa, esto no se ha tomado en consideración debido a la falta del material necesario para evaluar dichos factores. En el caso de las dimensiones referentes a la calidad y resolución de imagen, los resultados llevan a concluir que, para la creación del modelo principal, la nitidez de las imágenes es importante para el proceso de entrenamiento, ya que las acciones observadas en las imágenes se deben distinguir de manera clara y con la mejor calidad y resolución de imagen posible.
3. De forma general, respecto a la efectividad de los modelos entrenados, se determina que han pasado de manera satisfactoria las pruebas realizadas por los autores, obteniendo porcentajes de efectividad definitivamente altos como lo son un 96% y 100% para los modelos principal y secundario respectivamente. Además, se puede afirmar que la efectividad del modelo principal puede aumentar según la cantidad de pruebas realizadas por los autores.

VII. RECOMENDACIONES

A raíz de todo el proceso realizado en esta investigación, se tiene muy claro que existen aspectos que pueden tomarse en cuenta en futuras investigaciones, es por ello que como recomendación se plantea la evaluación de los factores externos en el ámbito de la iluminación, este factor sería de mucha influencia en los resultados de las investigaciones que tengan el mismo propósito de esta, ya que se demostró que el aumento de iluminación genera que el modelo entrenado no muestre los resultados que se esperan.

Como segunda recomendación, se propone realizar el reconocimiento de imágenes con múltiples acciones, debido a que en este estudio se ha definido una forma binaria de realizar el reconocimiento de gestos corporales, es decir, solo se puede diferenciar entre dos acciones y por ende obtener solo dos posibles resultados. A esta recomendación se le suma que los datos de entrenamiento aumenten de manera significativa de 2600 imágenes a 10000 aproximadamente, tomando en cuenta la cantidad de acciones que se plantean evaluar.

Otra recomendación para futuras investigaciones es el tema relacionado al aumento de resolución en las imágenes que se utilizan para realizar el proceso de entrenamiento, tal y como Fernández (2016) realiza la mejora en iluminación y aumento de resolución para la detección de rostros, ese principio se puede utilizar en futuros estudios para la obtención de mejores resultados, y de mayor efectividad por parte del modelo que se plantea entrenar.

REFERENCIAS

CANALES, P.L.G., 2017. Sistema biométrico para detección y reconocimiento de orejas basado en algoritmos de procesamiento de imágenes y redes neuronales profundas.

CASTILLO, C. y ANDRÉ, Y., 2015. Prototipo computacional para la detección y clasificación de expresiones faciales mediante la extracción de patrones locales binarios.

CASTILLO ESTRADA, C., 2014. Sistema de detección automático de movimientos corporales mediante técnicas de procesamiento digital de imágenes.

CUARESMA, T. y YUDIT, M., 2019. GESTIÓN DE RECURSOS HUMANOS Y COMPORTAMIENTO LABORAL DE LOS COLABORADORES DEL HOTEL JAZMIN, LINCE-2019.

FABIÁN, F. y MARIEL, K., 2016. Modelo para optimizar la detección de rostros en secuencias de video.

FUJIYOSHI, H., HIRAKAWA, T. y YAMASHITA, T., 2019. Deep learning-based image recognition for autonomous driving. IATSS Research, vol. 43, no. 4, pp. 244252. ISSN 0386-1112. DOI 10.1016/j.iatssr.2019.11.008.

GUERRERO, J., 2011. Técnicas de procesamiento de imágenes estereoscópicas. ,

JUAN, R. y CHACON, M., 2021. Redes neuronales artificiales para el procesamiento de imágenes, una revisión de la última década.

KAMIENKOWSKI, J.E., BENGOLEA MONZÓN, G., B. BIANCHI y FERRER, L., 2016. Estimadores computacionales de la predicción humana durante la lectura. Simposio Argentino de Inteligencia Artificial (ASAI 2016) - JAIIO 45 (Tres de Febrero, 2016).

MACAS, H.G.R., 2020. Algoritmo de reconocimiento de personas mediante procesamiento digital de imágenes usando MATLAB. Polo del Conocimiento: Revista científico - profesional, vol. 5, no. 8, pp. 1240-1252. ISSN 2550-682X.

MARCOS, A., PIZARRO, D., MARRÓN, M. y MAZO, M., 2013. Captura de movimiento y reconocimiento de actividades para múltiples personas mediante un enfoque bayesiano. Revista Iberoamericana de Automática e Informática Industrial RIAI, vol. 10, no. 2, pp. 170-177. ISSN 1697-7912. DOI 10.1016/j.riai.2013.03.007.

MASCHLER, B., KAMM, S., JAZDI, N. y WEYRICH, M., 2020. Distributed Cooperative Deep Transfer Learning for Industrial Image Recognition. Procedia CIRP, vol. 93, pp. 437-442. ISSN 2212-8271. DOI 10.1016/j.procir.2020.03.056.

MAZA, M.I.Z.S. de la, 2013. Metodología para el desarrollo de nuevas técnicas y algoritmos en el reconocimiento inteligente de imágenes submarinas.

MOREANO, J.A.C., PULLOQUINGA, R.H.M., LAGLA, G.A.F., CHISAG, J.C.C. y PICO, O.A.G., 2017. Reconocimiento facial con base en imágenes. Boletín Redipe, vol. 6, no. 5, pp. 143-151. ISSN 2256-1536.

MORENO, M. y DAVID, F., 2015. Reconocimiento de gestos corporales, utilizando procesamiento digital de imágenes para activar sistema de alarma.

NAIEMI, F., GHODS, V. y KHALESİ, H., 2021. A novel pipeline framework for multi oriented scene text image detection and recognition. Expert Systems with Applications, vol. 170, pp. 114549. ISSN 0957-4174. DOI 10.1016/j.eswa.2020.114549.

SANTOS, D., DALLOS, L. y GAONA-GARCÍA, P.A., 2020. Algoritmos de rastreo de movimiento utilizando técnicas de inteligencia artificial y machine learning. Información tecnológica, vol. 31, no. 3, pp. 23-38. ISSN 0718-0764. DOI 10.4067/S0718-07642020000300023.

SERRANO, S., 2018. Algoritmos de aprendizaje automático para análisis y predicción de datos. En: Accepted: 2018-12-05T22:13:54Z [en línea], [Consulta: 24 abril 2021]. ISSN 2072-568X. Disponible en: <http://www.redicces.org.sv:80/jspui/handle/10972/3626>.

SREELA, S.R. y IDICULA, S.M., 2018. Action Recognition in Still Images using

Residual Neural Network Features. *Procedia Computer Science*, vol. 143, pp. 563569. ISSN 1877-0509. DOI 10.1016/j.procs.2018.10.432.

SWARUP, P., CHEN, P., HOU, R., QUE, P., LIU, P. y KONG, A.W.K., 2021. Giant panda behaviour recognition using images. *Global Ecology and Conservation*, vol. 26, pp. e01510. ISSN 2351-9894. DOI 10.1016/j.gecco.2021.e01510.

TALAVERA, C. y JESUS, R., 2018. Modelo para la detección de rostros en secuencias de Video con Baja Calidad. En: Accepted: 2018-08-29T14:22:41Z, Universidad Nacional de San Agustín de Arequipa.

VILLA, J.S.V. y VELÁSQUEZ, M.A.V., 2012. Redes neuronales artificiales aplicadas al procesamiento de video. *REVISTA POLITÉCNICA*, vol. 8, no. 14, pp. 31-37. ISSN 1900-2351.

YOSHIDA, I. y HIROSHI, A., 2015. Diseño de un algoritmo de estabilización de video orientado a la detección de personas. En: Accepted: 2015-06-04T03:42:55Z, Pontificia Universidad Católica del Perú.

YUAN, Y., CHEN, L., WU, H. y LI, L., 2021. Advanced agricultural disease image recognition technologies: A review. *Information Processing in Agriculture*.

ZHANG, Q., ZHANG, X., MU, X., WANG, Z., TIAN, R., WANG, X. y LIU, X., 2021. Recyclable waste image recognition based on deep learning. *Resources, Conservation and Recycling*, vol. 171, pp. 105636. ISSN 0921-3449. DOI 10.1016/j.resconrec.2021.105636.

ANEXOS.

Tabla 5. Ficha de registro de la dimensión fondo de imagen.

Ficha de registro para fondo de imagen						
Investigadores	Cristhian Valentin Gallo Valladares Diana Estefany Nunura Garcia		Tipo de prueba	DESCRIPTIVO		
Institución	Servicios Generales y Transporte Fresco Pez S.R.L.					
Dimensión de estudio Procesamiento de imágenes	Fondo de imagen					
Fecha de Inicio	01/07/2021	Fecha final	15/10/2021			
Variable	Indicador	Medida	Fórmula			
Procesamiento de imagen	Posición de imagen Intensidad de color	Deciles y Cuartiles Escala RGB				
#	Fecha de revisión	Identificación del procesamiento	Tema (Imágenes del sistema de seguridad)	Contenido específico	Coordenadas	Intensidad de color
1	13-10-2021	foto129	Venta de producto	Pesando	X: 312; Y: 300	211,211,211
2	13-10-2021	foto131	Venta de producto	Pesando	X: 375; Y: 402	143,143,143
3	13-10-2021	foto137	Venta de producto	Pesando	X: 338; Y: 236	186,186,186
4	13-10-2021	foto333	Venta de producto	Cobrando	X: 332; Y: 435	62,62,62
5	13-10-2021	foto368	Venta de producto	Cobrando	X: 278; Y: 398	134,134,134
6	13-10-2021	foto1583	Venta de producto	Cobrando	X: 290; Y: 517	56,56,56
7	13-10-2021	foto1767	Venta de producto	Cobrando	X: 300; Y: 406	76,76,76
8	13-10-2021	imgs2765	Venta de producto	Cobrando	X: 258; Y: 482	73,73,73
9	13-10-2021	photo3212	Venta de producto	Cobrando	X: 216; Y: 472	41,41,41
10	13-10-2021	image324	Venta de producto	Pesando	X: 304; Y: 214	46,46,46
11	13-10-2021	foto6543	Venta de producto	Pesando	X: 305; Y: 235	248,248,248
12	13-10-2021	foto4653	Venta de producto	Pesando	X: 306; Y: 241	237,237,237
13	13-10-2021	imgs0	Venta de producto	Existencia de clientes	X: 147; Y: 212	62,62,62
14	13-10-2021	imgs9	Venta de producto	Existencia de clientes	X: 147; Y: 212	62,62,62
15	13-10-2021	imgs15	Venta de producto	Existencia de clientes	X: 147; Y: 212	62,62,62
16	13-10-2021	imgs21	Venta de producto	Existencia de clientes	X: 147; Y: 212	62,62,62
17	13-10-2021	imgs1118	Venta de producto	Ausencia de clientes	X: 147; Y: 212	87,87,87
18	13-10-2021	imgs1131	Venta de producto	Ausencia de clientes	X: 147; Y: 212	87,87,87
19	13-10-2021	imgs1133	Venta de producto	Ausencia de clientes	X: 147; Y: 212	87,87,87
20	13-10-2021	imgs1135	Venta de producto	Ausencia de clientes	X: 147; Y: 212	87,87,87

Tabla 6. Ficha de registro de la dimensión medición matricial.

Ficha de registro para medición matricial					
Investigadores		Cristhian Valentin Gallo Valladares Diana Estefany Nunura Garcia		Tipo de prueba	DESCRIPTIVO
Institución		Servicios Generales y Transporte Fresco Pez S.R.L.			
Dimensión de estudio Procesamiento de imágenes		Medición Matricial			
Fecha de Inicio		01/07/2021	Fecha final		15/10/2021
Variable		Indicador	Medida		Fórmula
Procesamiento de imagen		Nivel de iluminación	Lux (Lx)		Lumen / m ²
#	Fecha de revisión	Identificación del procesamiento	Tema (Imágenes del sistema de seguridad)	Contenido específico	Lux(lx)
1	13-10-2021	foto129	Venta de producto	Pesando	75 lx
2	13-10-2021	foto131	Venta de producto	Pesando	75 lx
3	13-10-2021	foto137	Venta de producto	Pesando	75 lx
4	13-10-2021	foto333	Venta de producto	Cobrando	75 lx
5	13-10-2021	foto368	Venta de producto	Cobrando	75 lx
6	13-10-2021	foto1583	Venta de producto	Cobrando	75 lx
7	13-10-2021	foto1767	Venta de producto	Cobrando	75 lx
8	13-10-2021	imgs2765	Venta de producto	Cobrando	75 lx
9	13-10-2021	photo3212	Venta de producto	Cobrando	75 lx
10	13-10-2021	image324	Venta de producto	Pesando	75 lx
11	13-10-2021	foto6543	Venta de producto	Pesando	75 lx
12	13-10-2021	foto4653	Venta de producto	Pesando	75 lx
13	13-10-2021	imgs0	Venta de producto	Existencia de clientes	75 lx
14	13-10-2021	imgs9	Venta de producto	Existencia de clientes	75 lx
15	13-10-2021	imgs15	Venta de producto	Existencia de clientes	75 lx
16	13-10-2021	imgs21	Venta de producto	Existencia de clientes	75 lx
17	13-10-2021	imgs1118	Venta de producto	Ausencia de clientes	75 lx
18	13-10-2021	imgs1131	Venta de producto	Ausencia de clientes	75 lx
19	13-10-2021	imgs1133	Venta de producto	Ausencia de clientes	75 lx
20	13-10-2021	imgs1135	Venta de producto	Ausencia de clientes	75 lx

Tabla 7. Ficha de registro de la dimensión calidad de imagen.

Ficha de registro para calidad de imagen					
Investigadores		Cristhian Valentin Gallo Valladares Diana Estefany Nunura Garcia		Tipo de prueba	
Institución		Servicios Generales y Transporte Fresco Pez S.R.L.			
Dimensión de estudio Procesamiento de imágenes		Calidad de imagen			
Fecha de Inicio		01/07/2021		Fecha final	
				15/10/2021	
Variable		Indicador		Medida	
				Fórmula	
Procesamiento de imagen		Cantidad de pixeles por pulgada		Pixeles (PX)	
#	Fecha de revisión	Identificación del procesamiento	Tema (Imágenes del sistema de seguridad)	Contenido específico	Número de pixeles
1	13-10-2021	foto129	Venta de producto	Pesando	128,44 ppp
2	13-10-2021	foto131	Venta de producto	Pesando	50,04 ppp
3	13-10-2021	foto137	Venta de producto	Pesando	128,44 ppp
4	13-10-2021	foto333	Venta de producto	Cobrando	128,44 ppp
5	13-10-2021	foto368	Venta de producto	Cobrando	105,32 ppp
6	13-10-2021	foto1583	Venta de producto	Cobrando	128,44 ppp
7	13-10-2021	foto1767	Venta de producto	Cobrando	128,44 ppp
8	13-10-2021	imgs2765	Venta de producto	Cobrando	128,44 ppp
9	13-10-2021	photo3212	Venta de producto	Cobrando	128,44 ppp
10	13-10-2021	image324	Venta de producto	Pesando	80,79 ppp
11	13-10-2021	foto6543	Venta de producto	Pesando	80,68 ppp
12	13-10-2021	foto4653	Venta de producto	Pesando	81,79 ppp
13	13-10-2021	imgs0	Venta de producto	Existencia de clientes	128,44 ppp
14	13-10-2021	lms9	Venta de producto	Existencia de clientes	128,44 ppp
15	13-10-2021	lms15	Venta de producto	Existencia de clientes	128,44 ppp
16	13-10-2021	lms21	Venta de producto	Existencia de clientes	128,44 ppp
17	13-10-2021	imgs1118	Venta de producto	Ausencia de clientes	128,44 ppp
18	13-10-2021	imgs1131	Venta de producto	Ausencia de clientes	128,44 ppp
19	13-10-2021	imgs1133	Venta de producto	Ausencia de clientes	128,44 ppp
20	13-10-2021	imgs1135	Venta de producto	Ausencia de clientes	128,44 ppp

Tabla 8. Ficha de registro de la dimensión resolución de imagen.

Ficha de registro						
Investigadores	Cristhian Valentin Gallo Valladares Diana Estefany Nunura Garcia			Tipo de prueba	DESCRIPTIVO	
Institución	Servicios Generales y Transporte Fresco Pez S.R.L.					
Dimensión de estudio Procesamiento de imágenes	Resolución de imagen					
Fecha de Inicio	01/07/2021	Fecha final	15/10/2021			
Variable	Indicador	Medida	Fórmula			
Procesamiento de imagen	Tamaño de imagen Nivel de pixeles	Píxeles (PX)			Tamaño de px -> (ppp * 2.54)/100	
#	Fecha de revisión	Identificación del procesamiento	Tema (Imágenes del sistema de seguridad)	Contenido específico	Cantidad de pixeles	Tamaño de pixeles
1	13-10-2021	foto129	Venta de producto	Pesando	1843200 px	3.26 mm
2	13-10-2021	foto131	Venta de producto	Pesando	277992 px	12.71 mm
3	13-10-2021	foto137	Venta de producto	Pesando	1843200 px	3.26 mm
4	13-10-2021	foto333	Venta de producto	Cobrando	1843200 px	3.26 mm
5	13-10-2021	foto368	Venta de producto	Cobrando	1185280 px	2.68 mm
6	13-10-2021	foto1583	Venta de producto	Cobrando	1843200 px	3.26 mm
7	13-10-2021	foto1767	Venta de producto	Cobrando	1843200 px	3.26 mm
8	13-10-2021	imgs2765	Venta de producto	Cobrando	1843200 px	3.26 mm
9	13-10-2021	photo3212	Venta de producto	Cobrando	1843200 px	3.26 mm
10	13-10-2021	image324	Venta de producto	Pesando	726138 px	2.05 mm
11	13-10-2021	foto6543	Venta de producto	Pesando	726291 px	2.04 mm
12	13-10-2021	foto4653	Venta de producto	Pesando	747355 px	2.08 mm
13	13-10-2021	imgs0	Venta de producto	Existencia de clientes	1843200 px	3.26 mm
14	13-10-2021	limgs9	Venta de producto	Existencia de clientes	1843200 px	3.26 mm
15	13-10-2021	limgs15	Venta de producto	Existencia de clientes	1843200 px	3.26 mm
16	13-10-2021	limgs21	Venta de producto	Existencia de clientes	1843200 px	3.26 mm
17	13-10-2021	imgs1118	Venta de producto	Ausencia de clientes	1843200 px	3.26 mm
18	13-10-2021	imgs1131	Venta de producto	Ausencia de clientes	1843200 px	3.26 mm
19	13-10-2021	imgs1133	Venta de producto	Ausencia de clientes	1843200 px	3.26 mm
20	13-10-2021	imgs1135	Venta de producto	Ausencia de clientes	1843200 px	3.26 mm

Tabla 9. Operacionalización de variable.

Título	Pregunta general	Objetivo general	Preguntas específicas	Objetivos específicos	Variable	Definición conceptual	Definición operacional	Dimensiones	Indicadores	Escala de medición
Implementación de un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes y video de los trabajadores de la empresa Servicios Generales y Transporte "Fresco Pez" Chulucanas – 2021	¿Cómo un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes se relaciona con las acciones de los trabajadores de la empresa Servicios Generales y Transporte "Fresco Pez" Chulucanas - 2021?	implementar un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes y video de los trabajadores de la empresa Servicios Generales y Transporte "Fresco Pez" Chulucanas – 2021	¿Cuál es la evaluación de los fondos de imagen en video por intensidad de color y posición?	Evaluar los fondos de imagen en video por intensidad de color y posición de imagen.	Procesamiento de imágenes	Se refiere a procesar las imágenes del mundo real de manera digital por medio de un computador. Es un tema muy amplio, en el que se incluyen estudios de física, matemáticas, ingeniería eléctrica, computación (Medina 2018).	Por cada dimensión de la variable se realizará una ficha de registro, además del uso de videos de seguridad, las imágenes serán analizadas en fases.	Fondo de imagen	Posición	Escala / Discreta
			¿Cuál es la medición matricial por imagen de video por iluminación en la empresa Servicios Generales y Transporte Fresco Pez S.R.L.?	Evaluar la medición matricial por el nivel de iluminación en la empresa Servicios Generales y Transporte Fresco Pez S.R.L.				Medición matricial	Nivel de iluminación	

			¿Cómo medir la calidad de las imágenes de video por cantidad de pixeles por	Medir la calidad de las imágenes de video por cantidad de pixeles por pulgada de				Calidad de imagen	Cantidad de pixeles por pulgada	
			pulgada de las cámaras de seguridad de la empresa Servicios Generales y Transporte Fresco Pez S.R.L.?	las cámaras de seguridad.						
			¿Cómo medir la resolución de las imágenes de video por el tamaño y nivel de pixeles de las cámaras de seguridad de la empresa Servicios Generales y Transporte Fresco Pez S.R.L.?	Medir la resolución de las imágenes de video por el tamaño y nivel de pixeles de las cámaras de seguridad de la empresa Servicios Generales y Transporte Fresco Pez S.R.L.				Resolución de imagen	Tamaño de imagen	Escala / Discreta
									Nivel de pixeles	

Anexo 1. Consentimiento informado.

DECLARACIÓN DE CONSENTIMIENTO INFORMADO

Por medio del presente documento confirmo mi consentimiento para participar en la investigación denominada: Implementación de un sistema para el reconocimiento de gestos corporales mediante el procesamiento de imágenes y video de los trabajadores de la empresa Servicios Generales y Transporte "Fresco Pez" Chulucanas – 2021

Se me ha explicado que mi participación consistirá en lo siguiente:

Entiendo que se utilizará material audiovisual ocultando la identidad de los trabajadores de mi empresa.

Se me ha explicado también que si decido participar en la investigación puedo retirarme en cualquier momento o no participar en una parte del estudio.

Acepto voluntariamente participar en esta investigación y comprendo qué cosas voy a hacer durante la misma.

Chulucanas, 09 de Julio de 2021

Nombre del participante:

Edwin Alberto Baca Chunga

DNI: 40002683



Investigador

Gallo Valladares Cristhian Valentin

DNI: 72224995



Investigadora

Nunura Garcia Diana Estefany

DNI: 74583165

Anexo 2. Código para obtener los fotogramas.

```
capturar.py parte_01.py parte_02.py
1 # capturar los fotogramas
2 import cv2
3
4 video_path = 'D:/usb/video.mp4' ##nombre del video
5 cap = cv2.VideoCapture(video_path)
6
7 img_index = 0
8
9 while (cap.isOpened()):
10     ret, frame = cap.read()
11     gray_effect = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
12     if ret == False:
13         break
14     cv2.imwrite('img1' + str(img_index) + '.png', gray_effect) # nombre de las imagenes
15     img_index += 1
16
17 cap.release()
18 cv2.destroyAllWindows()
```

Anexo 3. Directorio de imágenes

Este equipo > Datos (D:) > usb > dataset

Nombre	Fecha de modificación	Tipo
test	9/10/2021 11:25	Carpeta de archivos
train	8/10/2021 22:37	Carpeta de archivos
validation	8/10/2021 22:40	Carpeta de archivos

Anexo 4. Código para entrenar el modelo.

```
capturar.py  parte_01.py  parte_02.py
1  import os
2  import tensorflow as tf
3
4  from tensorflow.keras.models import Sequential
5  from tensorflow.keras.layers import Conv2D, MaxPool2D, Flatten, Dense
6  from tensorflow.keras.optimizers import RMSprop
7
8  from tensorflow.keras.preprocessing.image import ImageDataGenerator
9
10 def model_cnn():
11     model = Sequential()
12     model.add(Conv2D(filters=32, kernel_size=(3,3), activation='relu', input_shape=(150, 150, 3)))
13     model.add(MaxPool2D(2, 2))
14
15     model.add(Conv2D(filters=64, kernel_size=(3,3), activation='relu'))
16     model.add(MaxPool2D(2, 2))
17
18     model.add(Conv2D(filters=128, kernel_size=(3,3), activation='relu'))
19     model.add(MaxPool2D(2, 2))
20
21     model.add(Conv2D(filters=256, kernel_size=(3,3), activation='relu'))
22     model.add(MaxPool2D(2, 2))
23
24     model.add(Flatten())
25
26     model.add(Dense(512, activation='relu'))
27     model.add(Dense(1, activation='sigmoid'))
28
29     optimizador = RMSprop(lr=1e-4)
30     loss_function = 'binary_crossentropy'
31
32     model.compile(optimizer=optimizador, loss=loss_function, metrics=['acc'])
33     model.summary()
34     return model
35
36 def pre_process_image():
37
38     train_directorio = 'D:/usb/dataset/train/'
39     train_datagen = ImageDataGenerator(rescale=1./255)
40
41     train_generator = train_datagen.flow_from_directory(
42         train_directorio,
43         target_size=(150,150),
44         color_mode='rgb',
45         class_mode='binary',
46         batch_size=32)
47
48     validation_directorio = 'D:/usb/dataset/validation/'
49     validation_datagen = ImageDataGenerator(rescale=1./255)
50
51     validation_generator = validation_datagen.flow_from_directory(
52         validation_directorio,
53         target_size=(150,150),
54         color_mode='rgb',
55         class_mode='binary',
56         batch_size=32)
57
58     return train_generator, validation_generator
59
60 def train_model(model, train, validat):
61     model.fit_generator(train, steps_per_epoch=62, epochs=30, validation_data=validat, validation_steps=32)
62     model.save_weights('model_entrenado.h5')
63
64 nn_model = model_cnn()
65 train_data, valid_data = pre_process_image()
66 train_model(nn_model, train_data, valid_data)
```

Anexo 5. Código para reconocer gestos.

```
capturar.py  parte_01.py  parte_02.py
1  import tensorflow as tf
2  from keras.preprocessing.image import load_img, img_to_array
3  from parte_01 import model_cnn
4
5  image_path='D:/usb/dataset/test/foto1756.png'# imagen para determinar acción
6
7  img = load_img(image_path, target_size=(150,150))
8  img = img_to_array(img)
9  img = img.reshape(1, 150, 150, 3)
10 img = img.astype('float32')
11
12 model = model_cnn()
13 model.load_weights("model_entrenado.h5")
14
15 prediction = model.predict(img)
16
17 if prediction == 0:
18     print("Cobrando")
19 else:
20     print("Pesando")
```