



Universidad
Señor de Sipán

**FACULTAD DE INGENIERÍA, ARQUITECTURA
Y URBANISMO**

**ESCUELA PROFESIONAL DE INGENIERÍA DE
SISTEMAS**

TESIS

**Sistema de reconocimiento de rostros mediante
cámaras de seguridad interna**

**PARA OPTAR EL TÍTULO PROFESIONAL DE
INGENIERO DE SISTEMAS**

Autor (es)

Bach. Merino Ancajima Jhenson Jhampier

ORCID (<https://orcid.org/0000-0001-5256-0374>)

Asesor(a)

Mg. Mejia Cabrera Heber Ivan

ORCID (<https://orcid.org/0000-0002-0007-0928>)

Infraestructura, Tecnología y Medio Ambiente

Pimentel – Perú

2023

DETECTOR DE ROSTROS MEDIANTE CAMARAS DE SEGURIDAD INTERNA

Aprobación del jurado

MG. SAMILLAN AYALA ALBERTO ENRIQUE

Presidente del Jurado de Tesis

MG. ARCILA DIAZ JUAN CARLOS

Secretario del Jurado de Tesis

MG. MEJIA CABRERA HEBER IVAN

Vocal del Jurado de Tesis




DECLARACIÓN JURADA DE ORIGINALIDAD

Quien suscribe la **DECLARACIÓN JURADA**, soy estudiantes del Programa de Estudios de la Escuela de Ingeniería de Sistemas de la Universidad Señor de Sipán S.A.C, declaro bajo juramento que soy autor del trabajo titulado:

DETECCION DE ROSTROS MEDIANTE CAMARAS DE SEGURIDAD INTERNA

El texto de mi trabajo de investigación responde y respeta lo indicado en el Código de Ética del Comité Institucional de Ética en Investigación de la Universidad Señor de Sipán (CIEI USS) conforme a los principios y lineamientos detallados en dicho documento, en relación a las citas y referencias bibliográficas, respetando al derecho de propiedad intelectual, por lo cual informo que la investigación cumple con ser inédito, original y autentico.

En virtud de lo antes mencionado, firman:

Merino Ancajima Jhenson Jhampier	DNI: 70095518	
-------------------------------------	------------------	---

Pimentel, 31 de Marzo de 2023.

Dedicatoria

Dedico este proyecto de tesis a Dios y a mis abuelos. A Dios porque ha estado conmigo en cada paso que doy, cuidándome y dándome fortaleza para continuar.

A mis abuelos, quienes a lo largo de mi vida han velado por mi bienestar y educación siendo mi apoyo en todo momento depositando su entera confianza en cada reto que se me presentaba sin dudar ni un solo momento en mi inteligencia y capacidad. Es por ello que soy lo que soy ahora.

Agradecimientos

A cada miembro de mi familia, por siempre haberme dado su fuerza y apoyo incondicional que me han ayudado y llevado hasta donde estoy ahora.

A mi enamorada que siempre estuvo apoyándome en todo momento y siempre confió en mí.

Por último, a mis asesores que gracias a ellos pude desarrollar esta tesis.

Índice

Dedicatoria	IV
Agradecimientos.....	V
Índice de tablas y figuras.....	VIII
ÍNDICE DE TABLAS	VIII
Resumen.....	XI
Abstract.....	XII
I. INTRODUCCIÓN.....	13
1.1. Realidad problemática.....	13
Figura 1: Proceso de identificación de rostros. Fuente: Dang & Sharma	38
1.2. Formulación del problema	40
1.3. Hipótesis.....	40
1.4. Objetivos.....	41
1.5. Teorías relacionadas al tema.....	41
II. MATERIALES Y MÉTODO	50
1.1. Tipo y Diseño de Investigación	50
1.2. Variables, Operacionalización	50
1.3. Población de estudio, muestra, muestreo y criterios de selección ...	53
1.4. Técnicas e instrumentos de recolección de datos, validez y confiabilidad.....	53
1.5. Procedimiento de análisis de datos	53
1.6. Criterios éticos.....	54
III. RESULTADOS Y DISCUSIÓN	56
2.1. Resultados.....	56
2.2. Discusión	61
IV. CONCLUSIONES Y RECOMENDACIONES.....	64
3.1. Conclusiones	64

3.2. Recomendaciones	65
Referencias	66
ANEXOS	69

Índice de tablas y figuras

ÍNDICE DE TABLAS

TABLA I: Operacionalización de la variable independiente	51
TABLA II: Operacionalización de la variable dependiente	52
TABLA III: Escala Valorativa de cuantificación de indicadores (Variable Independiente).	
ÍNDICE DE TABLAS	VIII
57	
Tabla IV: Pruebas de Revocación	59
Tabla V: Resultado Valor F. Fuente: Elaboración propia.....	60

ÍNDICE DE FIGURAS

Figura 1: Proceso de	
Figura 1: Proceso de identificación de rostros. Fuente: Dang & Sharma	38
Figura 1: Proceso de identificación de rostros. Fuente: Dang & Sharma	38
Figura 2: Feature de ventanas del Método Haar Cascade	42
. Figura 3: Entrenamiento de imágenes positivas y negativas.....	44
Figura 4: Esquema general de reconocimiento de rostros en videos 2019	45
Figura 5. Arquitectura de red neuronal convolucional (CNN) y el proceso de entrenamiento	46
Figura 6. Red de neuronas.....	47
Figura 6: INDICADOR - PRECISIÒN.	57
Figura 7: Gráfico de Revocación	59
Figura 8: Gráfico de Revocación	60
Figura 9: Primera Prueba de detección de rostros	62
Figura 10: Segunda Prueba de detección de rostros	63
Figura 11: Interfas de entrenamiento	71
Figura 12: Archivo XML con la red neuronal entrenada	71
Figura 13: Imágenes Positivas	72
Figura 14: Imágenes Negativas.....	72
Figura 15: Esquema de la red neuronal	73
Figura 16: Paso 1 de la instalación de Python	74
Figura 18: Finalización de instalación de Python	75
Figura 19: Instalación de la librería OpenCv	76
Figura 20: Copiar línea de código imutils	77
Figura 21: instalación de la librería imutils.....	77
Figura 22: Copiar línea de código de MediaPipe.....	78

Figura 23: Instalación MediaPipe	78
Figura 24: Interfaz gráfica para la detección de rostros	83
Figura 25: Detección de rostros en imágenes fotográficas	84
Figura 26: Detección de rostros en imágenes de video	84
Figura 27: Data base de rostros detectados.....	85
38	
Figura 2: Feature de ventanas del Método Haar Cascade	42
. Figura 3: Entrenamiento de imágenes positivas y negativas.....	44
Figura 4: Esquema general de reconocimiento de rostros en videos 2019.....	45
Figura 5. Arquitectura de red neuronal convolucional (CNN) y el proceso de entrenamiento	46
Figura 6. Red de neuronas.....	47
Figura 6: INDICADOR - PRESICIÓN.	57
Figura 7: Gráfico de Revocación	59
Figura 8: Gráfico de Revocación	60
Figura 9: Primera Prueba de detección de rostros	62
Figura 10: Segunda Prueba de detección de rostros	63
Figura 11: Paso 1 de la instalación de Python	74
Figura 13: Finalización de instalación de Python	75
Figura 14: Instalación de la librería OpenCv	76
Figura 15: Copiar línea de código imutils	77
Figura 16: instalación de la librería imutils.....	77
Figura 17: Copiar línea de código de MediaPipe.....	78
Figura 18: Instalación MediaPipe	78
Figura 19: Interfaz gráfica para la detección de rostros	83
Figura 20: Detección de rostros en imágenes fotográficas	84
Figura 21: Detección de rostros en imágenes de video	84

Figura 22: Data base de rostros detectados..... 85

38

Resumen

La presente tesis realiza el análisis y evaluación de las diferentes técnicas y algoritmos de visión artificial para ambientes semi – controlados, en donde se utilizó opencv con python como herramienta de trabajo.

El principal problema que se presenta en esta investigación es que las características del rostro varían en posición o gestos con el paso del tiempo, calidad de la imagen, luminosidad, por lo que se torna un poco más complejo y por tal motivo no permite obtener resultados al 100%. Para esta investigación se pudo realizar los rostros en las personas dentro de un ámbito social en donde se detectarán las personas que ingresen o salgan de un local. La ejecución de este proyecto y validación durante las fases de prueba y operativas se considera necesaria y fundamental para poder avanzar en el conocimiento de los episodios de detección facial. Del mismo modo, se considera también básico para poder realizar tareas de evaluación y gestión de la calidad del reconocimiento facial.

Teniendo como objetivo principal el poder aplicar algoritmos de visión artificial para la detección del rostro de personas en cámaras de seguridad internas.

Para esta investigación utilizamos la técnica de HaarCascade junto con una Red Neuronal Convolucionante, las cuales permitieron la detección de rostros, las cuales comparándose con otras técnicas y algoritmos arrojaron resultados de un 98.1% permitiendo así poder seguir avanzando con el alto nivel de efectividad para la investigación realizada.

Palabras Clave: Detección de rostro, Algoritmo, OpenCv, Python.

Abstract

This thesis carries out the analysis and evaluation of the different artificial vision techniques and algorithms for semi-controlled environments, where opencv with python was used as a work tool.

The main problem that arises in this investigation is that the characteristics of the face vary in position or gestures over time, image quality, light, which makes it a little more complex which does not allow 100% results to be obtained. . For this investigation, it was possible to carry out the faces of people within a social environment where people who enter or leave a premises will be detected. The execution of this project and validation during the test and operational phases is considered necessary and fundamental to be able to advance in the knowledge of facial detection episodes. In the same way, it is also considered basic to be able to carry out tasks of evaluation and quality management of facial recognition.

With the main objective of being able to apply artificial vision algorithms for the detection of people's faces in internal security cameras.

For this research we used the haarcascade technique together with the convolutional neural network, which allowed the detection of faces, which, when compared with other techniques and algorithms, yielded results of 98.1%, thus allowing us to continue advancing with the high level of effectiveness for the research carried out.

Keywords: Face detection, Algorithm, OpenCv, Python.

I. INTRODUCCIÓN

1.1. Realidad problemática.

Shivalila Hangaragi [1], en su investigación denominada “Face Detection and Recognition Using Face Mesh and Deep Neural Network” de la Universidad de Amrita Vishwa Vidyapeetham, nos habla que la tecnología puede comparar rostros humanos a partir de una imagen y la base de datos almacenada de rostros. El modelo se extrae las características faciales de la imagen de entrada para autorizar a los usuarios a través de los servicios de verificación de identidad. Estos modelos se implementan diferentes entes públicos y privados.

Una aplicación principal del reconocimiento facial es la biometría facial. Es el más preferido, referente biométrico por su velocidad. También es fácil de implementar e implementar sin interacción humana. Es el método de autenticación y autorización de la identificación de la persona a partir de los rasgos faciales extraídos. Él captura, analiza y combina patrones en función de los rasgos faciales de la persona. Los pasos involucrados en el proceso son:

- Detección y localización de rostros humanos en la imagen
- Extracción de los rasgos faciales
- Comparación del rostro detectado con la base de datos.

Algoritmos como caras propias, histogramas de patrones binarios locales, caras de Fisher, función de escala invariable Transform y Speed Up Robust Features se utilizan para extraer la función. Enfrentar puntos de referencia como la distancia. Entre los dos ojos, la profundidad entre las orbitas de los ojos, el límite de la cara y la forma de los labios y el pómulo se utiliza para comparar la imagen de entrada de la consulta con las figuras del data base. El mayor beneficio del tratamiento facial sistemas de reconocimiento es que pueden identificarse a granel sin la intervención de la interacción humana. Hay muchas aplicaciones de tales sistemas. Algunas aplicaciones incluyen la instalación de sistemas de reconocimiento facial en lugares públicos como:

centros comerciales, aeropuertos y oficinas para biometría. Factores como la expresión, la iluminación, la pose y el ruido durante la imagen. La captura puede

alterar la productividad de los sistemas de reconocimiento facial.

La variación en la iluminación y la variación de la pose facial tienen un mayor impacto en la precisión de estos sistemas. La orientación de la iluminación conduce a cambios en la posición y forma de la sombra, cambios en los reflejos y también en la inversión de los gradientes de contraste. La orientación y alineación correctas de la pose de la cara conducen a que se extraen características que son robustas a las variaciones geométricas. El modelo propuesto supera los problemas encontrados en el sistema de reconocimiento facial y da resultados eficientes. El modelo usa Face mesh para identificar y detectar el rostro. Debido a Face mesh, el modelo funciona en una variedad de condiciones, como iluminación y fondo variables. El modelo también puede manejar imágenes no frontales del ser humano de diferentes años y razas. Imágenes del conjunto de datos de cara salvaje etiquetada (LWF) capturados en tiempo real se utilizan para la cognición del modelo. Durante la prueba, si los puntos de referencia de la cara de la imagen de prueba coinciden con los puntos precisos de la cara de cualquiera de las imágenes de entrenamiento, el modelo proporciona el nombre de la otra persona que genera el modelo como "desconocido". El modelo propuesto logra una precisión del 94,23% para el reconocimiento facial.

La detección de rostro es donde refiere a los mecanismos de computadoras en el cual es factible comprobar si se encontró o no una cara dentro de una figura digital brindada, donde se podrá verificar su posición dentro de la imagen, su tamaño. El poder detectar objetos es una tarea que se estudia y se aplica en la rama de visión computacional. Para poder detectar un rostro específico, se busca en la figura las particularidades que se observa que son relacionado a él, sin embargo, en la medida que se quiere detectar un rostro se aumenta la dificultad ya que los rostros se encuentran en diferentes poses, con bastantes particularidades visuales que cambian su aspecto; todo lo que coopera a nuestra mente a diferenciar un ser humano de otro, se agrega más punto de dificultad cuando se trata de localizar la parte facial mediante el procesamiento de imágenes.

Ming-Hsuan, Kriegman [2], en su investigación "Detecting Faces in Images: A

Survey", nos menciona los aspectos que se encadenan de manera directa con los desafíos de la detección de rostros:

La variación de las materias estructurales: Las discrepancias entre rasgos faciales contiene rasgos propios de la cara, como la silueta, tonalidad y magnitud, tal como los cambios en la ausencia o presencia de rasgos más específicos, como por ejemplo los diferentes aspectos o artículos que posee un rostro.

Posición del rostro: La postura en la que se encuentra una cara ante el dispositivo digital afecta de manera directa en el aspecto de este, la posición podría dañar la visión de partículas fundamentales en los detalles faciales.

Rasgos en el Rostro: Perjudica de manera rápida y concreta el aspecto de los rasgos faciales.

Oclusión: Refiere a la probabilidad de hallar, en la figura de entrada, rasgos faciales inicialmente escondidos; que podría ser por otros objetos o por otros rasgos faciales.

Direccionamiento de la figura: La ubicación del dispositivo digital con dirección a la cara afecta la manera en que se observa la cara.

Condiciones de la figura: Agrupa todos los detalles que podrían perjudicar la fase donde se adquiere la figura, ya sea los componentes de la cámara, la luz, el movimiento.

Existen muchos problemas estrechamente relacionados con la detección de rostros. La localización facial tiene como objetivo determinar la imagen.

posición de un solo rostro; esta es una detección simplificada del problema con la suposición de que una imagen de entrada contiene un solo rostro. El objetivo de la detección de rasgos faciales es para detectar la presencia y ubicación de características, como ojos, nariz, fosas nasales, cejas, boca, labios, oídos, etc., con la suposición de que solo existe un rostro en una imagen.

El reconocimiento facial o la identificación facial compara una imagen de entrada (sonda) contra una base de datos (galería) e informa una coincidencia.

El propósito de la autenticación facial es verificar el reclamo de la identidad de un individuo en una entrada imagen, mientras que los métodos de seguimiento de rostros continuamente estiman la ubicación y posiblemente la orientación de un rasgo facial en una secuencia de figuras en transmitiéndose de manera directa. El

reconocimiento de expresiones se refiere a la identificación de los estados afectivos (feliz, triste, disgustado, etc.) de los humanos.

Evidentemente, la detección de rostros es el primer paso en cualquier proceso automatizado.

Vale la pena mencionar que muchos documentos usan el término "detección de rostros", pero los métodos y los resultados experimentales sólo muestran que un solo rostro se localiza en una imagen de entrada. En este artículo, diferenciamos la captura de rasgos faciales de la localización de los mismos ya que el segundo es un problema simplificado del primero.

Mientras tanto, nos centramos en los métodos de detección de rostros en lugar de métodos de seguimiento. Entre los métodos de detección de rostros, los basados en los algoritmos de aprendizaje han llamado mucho la atención recientemente y han demostrado excelentes resultados. Dado que estos métodos basados en datos dependen en gran medida de los conjuntos de entrenamiento, también discuten varias bases de datos adecuadas para esta tarea.

Un problema importante es cómo evaluar el desempeño de los métodos de detección propuestos. Muchos artículos recientes de detección de rostros comparan el desempeño de varios métodos, generalmente en términos de detección y falsa alarma.

También vale la pena notar que muchas métricas han sido adoptadas para evaluar algoritmos, como el tiempo de aprendizaje, el tiempo de ejecución, el número de muestras requeridas en el entrenamiento, y la relación entre las tasas de detección y las falsas alarmas.

La evaluación se vuelve más difícil cuando los investigadores usan diferentes definiciones para tasas de detección y falsas alarmas. En este documento, la tasa de detección se aclara como la unión del número de rostros detectados correctamente y del número de rostros determinada por un ser humano. Una región de la imagen identificada como cara por un clasificador se considera correctamente detectada si la región de la imagen cubre más de un cierto porcentaje de una cara en la imagen. En general, los detectores pueden cometer dos tipos de errores: falsos negativos en qué caras se pasan por alto, lo que da como resultado bajas tasas de detección y falsos positivos en los que se declara que una región de la imagen es cara, pero no lo es. Una evaluación justa debe tomar estos factores en

consideración ya que uno puede ajustar los parámetros del método de uno para aumentar las tasas de detección mientras también aumentando el número de detecciones falsas. En este papel, nosotros discutir los conjuntos de datos de evaluación comparativa y los problemas relacionados en una evaluación justa.

Con más de 150 enfoques informados para la localización de rasgos faciales, la investigación en localización de caras tiene implicaciones más amplias para investigación de visibilidad computacional en la localización de objetos. Casi todos enfoques basados en modelos o basados en la apariencia para objetos 3D

reconocimiento se han limitado a objetos rígidos mientras que intentar realizar una identificación robusta en una amplia variedad de ubicaciones de cámaras y condiciones de iluminación. La localización de una cara puede verse como una dificultad de localización de dos clases en el que una región de la imagen se clasifica como una "cara" o "sin rostro". En consecuencia, la detección de rostros es una de las pocas intenta reconocer a partir de imágenes (no representaciones abstractas) una clase de objetos para los que hay una gran cantidad de variabilidad dentro de la clase (descrita anteriormente). también es uno de las pocas clases de objetos para los que esta variabilidad ha sido capturada usando grandes conjuntos de entrenamiento de imágenes y, por lo tanto, algunas de las técnicas de detección pueden ser aplicables a una clase mucho más amplia de problemas de reconocimiento.

La detección de rostros también ofrece desafíos interesantes para la clasificación de patrones subyacentes y técnicas de aprendizaje.

Cuando una imagen sin procesar o filtrada se considera como entrada a un clasificador de patrones, la dimensión del espacio de características es extremadamente grande (es decir, el número de píxeles en normalizado imágenes de entrenamiento). Las clases de imágenes faciales y no faciales son decididamente caracterizadas por funciones de distribución multimodal y es probable que los límites de decisión efectivos no sean lineales en el espacio de la imagen. Para ser efectivos, los clasificadores deben ser capaz de extrapolar a partir de un número modesto de entrenamiento muestras o ser eficiente cuando se trata de un gran número de estas muestras de entrenamiento de alta dimensión.

Con el objetivo de ofrecer un estudio completo y crítico de los métodos actuales de detección de rostros, esta información se organiza de la siguiente manera: En la

Parte dos, damos la revisión detallada de técnicas para detectar rostros en una sola imagen. evaluación comparativa los data set y los puntos de evaluación se analizan en la Sección 3.

Concluimos este artículo con una discusión de varias direcciones prometedoras para la detección de rostros.

Aunque informamos tasas de error para cada método cuando disponibles, las pruebas a menudo se realizan en conjuntos de datos únicos y, por lo tanto, las comparaciones son a menudo difíciles. Indicamos esos métodos que han sido evaluados con un conjunto de prueba disponible públicamente. Él se puede suponer que se utilizó un conjunto de datos único si no indicar el nombre del equipo de prueba

Hernández [3] , en su investigación “Desarrollo de un Nuevo Algoritmo de Detección Facial y su Aplicación a un Atril Automatizado” nos habla sobre la detección facial la cual consiste en analizar distintos tipos de fotografías e identificar dónde existen rostros, sin tomar importancia de a quién pertenece el rostro. Existen diversas técnicas que permiten la detección del rostro a través de distintos factores, como la separación de los ojos, distancia entre las fosas nasales y los labios. Se han desarrollado distintos tipos de algoritmos que, en porcentajes elevados, reconocen satisfactoriamente los rostros en una fotografía o video y a partir de esto se realiza el reconocimiento facial para sistemas de seguridad en el hogar u oficinas. Actualmente la detección facial forma parte del reconocimiento facial en dispositivos móviles y sistemas de seguridad en el hogar. El proceso de reconocimiento facial no consigue en todas las ocasiones un resultado satisfactorio porque depende de factores físicos de la obtención de la imagen como son: los cambios de luz que recibe la cara del individuo, el ángulo en que la fotografía fue tomada y las distintas resoluciones de la cámara empleada. En el artículo Comparative Testing of Face Detection Algorithms se menciona que las tareas que requieren del uso de la detección facial son requeridas más frecuentemente en la vida diaria de las personas (Degtyarev y Seredin, 2010). Por otro lado, en el trabajo de investigación Sistema de reconocimiento facial y realidad aumentada para dispositivos móviles se asegura que el hardware y software para sistemas de seguridad han tenido un gran impulso en los últimos años y el reconocimiento facial destaca entre estos

porque el uso de cámaras de video también ha aumentado en nuestro hogar, así como en el trabajo y lugares de recreación. El mejoramiento de los algoritmos para la detección facial traería grandes beneficios a todo tipo de usuarios, desde el mejor posicionamiento de un filtro digital usado en las redes sociales, hasta la integración de sistemas de seguridad en hogares particulares.

González [4], en su investigación “Implementación y Comparación de dos Algoritmos Supervisados en Redes Neuronales Convolucionales Orientadas a la Detección de Rostros para Ejecutarse en Hardware de Bajos Recursos” explica que, la inteligencia artificial se define de diferentes maneras con el transcurso del tiempo, pero la más recibida tiene cavidad de un equipo computacional que actúa de manera inteligente, se centra en el pensamiento de la astucia del ser humano, fundamentalmente en la faceta realizada donde se llega a un punto de vista; de esta manera, intenta igualar la manera donde las personas interpretan la noticia y a través de eso toma decisiones. A través de un equipo de cómputo, esto se logra al conseguir datos externos que sirven de fondo en la fase de aprender y de esta forma, estos aprendizajes, completan los objetivos o metas exactas y se adaptan de forma constante, superando este inicio podemos hallar varias fases que se despliegan de esta rama, esto se enfatiza en la vista computacional o vista sintética, que es un grupo de herramientas y metodologías que ayudan a conseguir, dar curso y verificar las figuras reales y son vistas y entendidas por un computador. Se puede asumir que la inteligencia artificial la es capaz de actuar de manera inteligente, así como las personas usan la visión y su capacidad de razonar para entender el universo en el que se encuentran, la visión artificial realiza igual resultado para conseguir que los computadores tengan la capacidad de recibir y entender una figura o secuencia de las mismas que se asocia a una particularidad anteriormente definida. Esto no da bastantes probabilidades a solucionar inconvenientes al llegar a mecanizar fases que necesitan de un humano para realizar una labor, fundamentalmente en cuestiones vinculados a la visión que se aplican de una manera más infalible con la utilización de la visión computacional, por tal motivo al implementar dicha tecnología se avanza bastante para diferentes ramas que requieran este tipo de herramientas.

C. Alcantara. *et al.* [5] elaboraron un trabajo denominado “Detección efectiva de rostros en imágenes utilizando descriptores basados en HOG”, en Santiago de Queretaro, Mexico. Este artículo propone la detección de rostros mediante descriptores basados en HOG, donde la detección de rostros se refiere a mecanismos computacionales que permite detectar si existe una cara en una figura digital obtenida, y por lo tanto detectar la pose y su tamaño. Para localizar un determinado objeto, se busca en la figura las particularidades visuales que son relacionados a él, pero, cuando se requiere ubicar o detectar un rasgo facial la complejidad se agranda ya que los rasgos faciales pueden tener un nivel de posiciones, con diferentes actos y particularidades visuales que varían su apreciación; toda particularidad que a simple vista apoya a nuestra mente a diferenciar a las personas, para la detección de rostros mediante imágenes digitales resulta más complejo. Para esta investigación se enumera los factores que se relacionan directamente con la localización de una cara:

La variación de las materias estructurales: Las discrepancias entre rasgos faciales contiene rasgos propios de la cara, como la silueta, tonalidad y magnitud, tal como los cambios en la ausencia o presencia de rasgos más específicos, como por ejemplo los diferentes aspectos o artículos que posee un rostro.

Posición del rostro: La postura en la que se encuentra una cara ante el dispositivo digital afecta de manera directa en el aspecto de este, la posición podría dañar la visión de partículas fundamentales en los detalles faciales.

Rasgos en el Rostro: Perjudica de manera rápida y concreta el aspecto de los rasgos faciales.

Oclusión: Refiere a la probabilidad de hallar, en la figura de entrada, rasgos faciales inicialmente escondidos; que podría ser por otros objetos o por otros rasgos faciales.

Direccionamiento de la figura: La ubicación del dispositivo digital con dirección a la cara afecta la manera en que se observa la cara.

Condiciones de la figura: Agrupa todos los detalles que podrían perjudicar la fase donde se adquiere la figura, ya sea los componentes de la cámara, la luz, el movimiento.

Capcha Huaman Alexander, Huaman Chile, José Luis [6], en su investigación sobre “Sistema de Reconocimiento Facial Basado En Los Algoritmos Haar Cascade,

Deep Faces y Luxand FacesSDK” Tienen como proposito detectar los efectos que tienen en sus capacidades de periodo, exactitud y recursos. A través de los mecanismos que se plantean. La investigación propuesta es aplicada, de un entorno cuantitativo y de diseño pre-experimental. La variable de causa de la utilización del aplicativo de reconocimiento de rostros se basa en la utilización de algoritmos Haar Cascade, DeepFace y Luxand FaceSDK. Para este proyecto se utilizó una data de 50 humanos y una muestra de 20 humanos de distintos aspectos ya que se apartó por conveniencia, y los mecanismos que se aplicaron fueron utilizados como un registro que administre tareas propias de las diferentes versiones de los Sistemas Operativos Windows con el propósito de calcular los indicativos. Los efectos son muy buenos, ya que se mejoró en los diferentes indicativos, el porcentaje de significancia fue de 0.01 y es menos a 0.05. De esta manera, se recomienda aplicar próximos aplicativos de investigación cuyas funciones sean de detecciones de movimiento para clases en línea.

Zequiel Ángel Jeremías Ambrogio [7], en su investigación presentó en su investigación denominada “Reconocimiento de objetos a través de la Metodología Haar Cascade” en Córdoba, Argentina. Presenta un tema que aborda el reconocimiento de objetos en figuras, un tema que se estudia desde varios años, pero tuvo su auge en temas a su aplicación y uso por los hallazgos que han sucedido en trabajos de visión por computadora durante este último tiempo. Es importante el tema y el poder que posee, se observa entre otras cosas, la gran aleatoriedad de sistemas y probabilidades que nos proporciona su aplicación, desde temas de salud, seguridad, controles biométricos, controles de tránsito entre otros. Cabe recalcar una diferencia significativa, no es igual una figura o una secuencia de figuras que un objeto, normalmente se habla con alguna ligeros de “detección de figuras” y no es la mejor manera de nombrarlo, porque una figura tiene diferentes objetos que la forman, por lo que la manera más adecuada de mencionar es “detección de objetos”; ya sea en una figura o en una secuencia de las mismas (como en partes de un video). En ese sentido, sería más adecuado conversar de estudios y procesamiento de figuras o sobre ellas. En este caso, se percibe como algo nuevo, como se señala, la detección de objetos y el procesamiento de figuras es un caso de estudio que viene sucediendo con anterioridad. Pero, las tácticas y

algoritmos que se utilizan para la detección de objetos aún son objeto de estudio, así como el ejemplo del matiz de piel híbrido (RGB-CbCrCg) que se utiliza para seleccionar píxeles de piel y no piel, o la metodología de avance diferencial de gran cantidad de objetos que se utilizan para extraer de una parte de la piel. Actualmente el poder de computador que disponemos ha sufrido mejoras sustancialmente en la cavidad que poseemos de lograr hacer análisis matemáticos avanzados y realizar algoritmos complicados, reduciendo espacios de los cálculos y temas que requiera. Como modelos de estos temas pueden funcionar cuando se utilizan para la localización de bordes, matices, magnitud, formas, escala de grises, diferentes tipos de filtros, gradiente de figuras, umbrales y otros. Particularmente se estudiará en esta investigación una de las técnicas y mecanismos libres para poder reconocer y localizar objeto/s dentro de una figura o frame; abarcando el tema de un caso práctico que se trata de la detección de automóviles en cámaras de control de tránsito de una ciudad de la provincia de Córdoba. Esencialmente en este documento, como primera parte, haremos una pequeña referencia al reconocimiento de objetos usando Haar Cascades; seguidamente, se propondrá una prueba ejemplo sobre una experiencia que se espera pueda ser el inicio para un reconocimiento y control de imágenes de poco recurso financiero, aquello se atiende a las potencias y probabilidades de la aplicación en diferentes ambientes de esta herramienta u otras similares.

González [8] en su tema "Sistema de Reconocimiento Facial con Deep Learning", en Leganés, Ecuador, nos mencionan que, en las últimas décadas, la inteligencia artificial (IA) aumentado a una manera bastante acelerada. La perseverante evolución de la ciencia aplicada en compañía con la gran capacidad de cómputo de los dispositivos, han dado acceso a que se realicen novedosas líneas de estudio que anteriormente no eran factibles en este rubro. Dentro del mercado IT, el rubro de la IA es el que reemplaza la categoría de mayor auge, pues las predicciones señalan que se elevaran a los 35 billones de dólares americanos (35.000 millones) para el año 2025.

Uno de los campos con más aumento dentro de la Inteligencia Artificial es la visión por computadora, ciencia que ha tenido un aumento también potente a través de

los últimos cinco años y que se prevé que seguirá aumentando con esta novedad a través de los siguientes años. “Mediante un nuevo documento de Practica, el mercado mundial de hardware y software de visión por computadoras seguirá en aumento hasta bordear de \$ 6,6 mil millones en 2015 a \$ 48,6 mil millones al año para el 2022.

Tan grande es la magnitud del crecimiento de la visión por computadora en las empresas, que es muy factible hallar sistemas que la componen actualmente. Ciertas aplicaciones de la visión por computadoras se brindan en los ambientes mencionados a continuación:

Militares: Los ejércitos de hoy en la actualidad se ayudan a través de técnicas de inteligencia artificial para lograr localizar de ejércitos enemigos en ciertas áreas. También, esta tecnología les permite avanzar en escoger metas previas al enviar de misiles guiados.

Industrias: Para la localización de productos en mal estado o localización, armado y supervisión de envolturas en fases logísticas. Otro ejemplo es del uso real se pueden hallar en Osperity (anteriormente se conocía como Osprey Informatics), empresa que trabaja con la visión por computadoras para la supervisión de pozos petroleros de forma virtual.

Salud: Un extenso porcentaje de la información médica se inician en figuras (radiografía, resonancia magnética, TACs, etc.). Por esto, han salido muchas tecnologías que utilizan la visión por computadoras para localizar anomalías en estos casos de experimentos de diagnósticos. Un echo verídico es el de Gauss Surgical, que emplea su ciencia aplicada llamada Triton para operaciones, que “capta escaners veloces de esponjas y depósitos que conllevan sangre y utiliza visión por computadoras. Ingresos visión artificial por mercado. Mercados globales: 2015-2022 por computadoras y aprendizajes automáticos para estimar cuanta sangre se llega a perder con una exactitud significativamente superiores” que los profesionales médicos.

Autos / Motores: Procedimiento de visión por computadoras que nos dan acceso a implementar cada vez con gran exactitud los sistemas de conducción autónoma de

autos. Se utilizan algoritmos que nos dan acceso a la localización de señales, seguir líneas y marcas en la pista, localización de peatones, etc.

Colaboración de visión virtual: La gran mayoría de personas está aceptando una forma cada vez más normal el uso de dispositivos que son fáciles de interactuar con electrodomésticos y otros dispositivos de uso diario.

Actualmente, los dispositivos como televisores de gama alta (smart TVs), refrigeradoras y celulares son utilizados a través gestos y/o audio. Como se puede mirar, la visión por computadoras se aplica a infinidad de casos, pertenecen a la mayoría de áreas de ventas. En tal sentido, esta investigación se enfocará en el área vinculada con el reconocimiento de rasgos faciales para estudiar las virtudes del machine learning en este aspecto. La motivación fundamental de esta investigación es trabajar un sistema de reconocimiento facial utilizando tácticas de deep learning que, cuando se implemente y se realicen los experimentos pertinentes, pueda ayudar a solucionar preguntas como: ¿es este rubro de la inteligencia artificial es lo más maduro como para localizar caras de personas de manera sencilla? ¿es capaz de localizar patrones e identificar características teniendo un conjunto de información extensa? ¿logrará seleccionar caras dependientemente del objetivo para el cual se ha entrenado el sistema (identificación de género, raza, años, expresiones faciales)?

García [9], en su tema “Integración y evaluación de sistemas de reidentificación de caras usando MTCNN y FaceNet en C++” en Sevilla, España.

Nos habla que, en las últimas temporadas, la localización de rasgos faciales se ha vuelto un tema de suma relevancia ya que los muchos sistemas que posee. Entre las que están aquellas vinculadas con el mando de accesibilidad a lugares o automóviles; el permitir o negar el acceso a tecnologías, tales como celulares inteligentes o computadoras; prevenir la estafa, principalmente en lo que se trata a la banca móvil; prevenir de bombardeos virtuales; o de video vigilancia avanzada. El acto que aumenta el interés de ciertas ramas en ella ayuda a que exista una mayor capacidad de uniones de investigación que buscan arreglos con mas precisión a diferentes inconvenientes. Hace muchos años, estos aplicativos se están utilizando en gran cantidad. Por ejemplo, en China diagnostica que la

capacidad de cámaras que se destinan a esto llegue a las 40000. Actualmente, en este país se utilizan para vigilar grupos étnicos pequeños, la localización de maleantes y el control de alumnos en los colegios. En 2017 la plataforma Alipay, de la compañía Alibaba, envió el servicio Smile To Pay que brinda el pago a través de tácticas de reconocimiento de expresiones faciales. En las Olimpiadas de Tokio de 2020 se implementaron sistemas con un procesador Intel Core i5 para la localizar de más de 300000 humanos, incluidos deportistas y personas voluntarias. Esto da hincapié a un mayor número de opositores, que comentan que, aunque estos sistemas puedan tener una mayor capacidad de ventajas, también brindan que se pierda la privacidad de las personas involucradas. También, se piensa que también podrían ocasionar un mayor caso de delitos como son acoso.

Granja, Moreno, Cabrera, & Valle [10] presentaron el documento denominado “Procesamiento de imágenes para la identificación de personas como sistema de seguridad en zonas domiciliarias” en Riobamba, Ecuador. En su investigación nos habla que la tecnología de reconocimiento facial últimamente, ha tenido gran salto por las aplicaciones y usos en lo que se trata a asuntos de seguridad, aunque, las investigaciones iniciales se dan en la época de los 60 y que después de 50 años se han mejorado sus métodos matemáticos y dispone de tecnología más sofisticadas, se considera esta ciencia aplicada como tema esencial en sistemas biométricos, se utilizan en diferentes instituciones y empresas privadas. La creación del actual tema de investigación se orienta al análisis de técnicas y algoritmos de reconocimiento de rasgos faciales para identificar personas e implementar dos algoritmos capaces de hacer el reconocimiento de rasgos faciales para identificar a un ser humano que está ingresada en un dataset, para esto utilizamos una base de datos de rostros públicos que contienen diferentes imágenes que fingen capacidades verdaderas para ser procesadas, y otro dataset de caras de domicilio para comprobar si es eficiente en condiciones reales de vigilancia en una casa. Cuando se realizó la incorporación de dos tácticas distintas de reconocimiento de expresiones faciales como metodología de seguridad casera se quiere encontrar la eficacia de identificación de cada táctica y se establecen requisitos mínimos de solicitudes para implementar el sistema de vigilancia domiciliaria, con probabilidades de exactitud altas en estos ambientes.

Chaves [11] presentaron su proyecto denominado “Diseño de un sistema de apoyo en seguridad basado en reconocimiento de múltiples rostros e identificación de usuarios” en Bogotá, Colombia. Habla sobre los grandes avances tecnológicos a nivel de hardware y software logrados en las últimas décadas han conseguido que actualmente existan máquinas con gran capacidad de procesamiento y que cada día toman mayor acción en la solución de problemas con aplicaciones en diversos campos. Las técnicas de aprendizaje de máquina permiten mediante el procesamiento de datos y el análisis de los mismos, construir modelos que resuelven tareas cada vez más complejas. Tareas como la detección objetos en imágenes o videos, o la predicción de una etiqueta para clasificar datos acordes con algunas características relevantes son solo un inicio de todo lo que puede lograr el aprendizaje automático. De esta forma, áreas de visiones computacionales, procesar la comunicación fluida o traducción automática, resuelven ciertas tareas empleando algoritmos de machine learning como una red neuronal o herramientas de aguate vectorial, entre otras técnicas. Puntualmente, el reconocimiento facial ha tenido un alto impacto en los sistemas de seguridad modernos, en gran parte por el aumento de los circuitos de cámaras instalados que permiten monitorear grandes áreas de interés en tiempo real. Ya existen programas capaces de realizar seguimiento por diferentes cámaras (tracking) e incluso realizar una construcción 3D de la cabeza de una persona a partir de una secuencia de imágenes. Por lo anterior, la meta de este trabajo de investigación es solucionar el inconveniente de reconocimiento de caras implementando una herramienta basada en aprendizaje automático para que realice la detección y clasificación de rostros sobre una imagen o vídeo en el contexto del problema. En el presente documento se consigna el desarrollo de cada una de los programas que realizan las tareas requeridas, así como el análisis de la sintonización de los parámetros elegidos para cada uno de los modelos empleados. Finalmente, se discuten los resultados obtenidos en la validación.

Yoanna Martínez Díaz [12] presentaron su trabajo llamado “Nuevos métodos para el reconocimiento automático de rostros en video” en Cuba. El reconocimiento de

rostros en videos tiene mucha importancia en diferentes aplicaciones como la video protección, la vigilancia y el control de ingresos. Requiere de distintas fases que se aplican a través del ingreso al sistema de un seguimiento de imágenes o frame que poseen las caras de las personas a las que se quiere identificar. En primer lugar, se determina, a través de la detección y el seguimiento, donde están ubicados los rostros en cada uno de los frame del video; después se pasa a la exhibición de las secuencias, y por último, a través de la comparación, obtenemos una respuesta de la identificación/verificación de los rostros que se analizaron.

La gran parte de la metodología para la identificación de rostros en videos que se proponen en la literatura se basan en conseguir elevados grados de eficacia en la fase de identificación/verificación, dejando a un lado el presupuesto computacional y suponiendo que las caras se detectaron y localizaron anteriormente. Para aplicaciones prácticas se necesita tener algoritmos que sean buenos y también eficientes, que puedan procesar una gran parte posible de flujos provenientes imágenes. Actualmente, los mejores resultados en cada fase se han obtenido utilizando métodos que se basan en redes neuronales profundas, pero se presentan varias limitaciones para utilizarlos, como el presupuesto computacional. Para detectar rostros en video, la metodología basada en técnicas para imágenes estáticas son las que más se utilizan. Entre ellos, la metodología de detección que desarrolló Viola y Jones es muy conocida y transmitida, porque permite detectar rostros con una muy baja calidad computacional. Aun así, se le presta poco interés a la unión de la información pasajera y espacial de la apariencia facial para detectar caras en videos cuando se ve que las caras tienen una forma con una repartición espacial que se obtiene al mantener con el tiempo. cuando se sigue, los enfoques que se basan en el seguimiento a través de la detección se consiguen ser de los más adecuados, porque trata con escenas de nivel dificultoso y son más grandes ante distintas modificaciones. La propuesta consiste en adaptar un detector en cada frame del video y después, juntar las detecciones que se obtuvieron para diseñar la trayectoria que corresponde a cada cara. Ya que el detector que existe todavía se encuentra alejado de la perfección, los criterios de respuesta para elegir un resultado de detección a una trayectoria determinan el éxito de la fase de seguimiento. Para representar las secuencias de rostros, se utilizan distintos tipos

de características. A pesar de que algunas han obtenido grandes valores de eficacia, gran parte de ellas son caras. Recientemente, descriptores binarios se han propuesto para distintas aplicaciones con el fin de reducir el presupuesto computacional y almacenamiento. Este tipo de rasgos brinda una solución gustosa para aplicaciones prácticas, donde se requiere un procesamiento veloz de altos volúmenes de datos y donde los recursos digitales tengan un tope. Para calcular el desempeño de la metodología de reconocimiento, un factor principal es el diseño de protocolos de evaluación apropiados, definidos en el data set de referencia. No obstante, se han publicado pocos data set que tengan protocolos que puedan conseguir los requisitos de los lugares reales y estallar todos los datos disponibles. Es por ello que, de vez en cuando, el rendimiento de los métodos en la actualidad en data set de referencia parece estar lleno. En este caso, los esfuerzos deben ser en medio del diseño de nuevos protocolos y no en la recolección de nueva información, que es más cara y menos recomendable. Sobre lo expuesto con anterioridad y los problemas presentados en la literatura relacionada, en esta investigación se trabajaron nuevas metodologías que apoyan a superar la eficiencia de la localización de rostros en aplicaciones de video protección. Se expande el método de detección de Viola y Jones al dominio espacio temporal, conservando sus grados de eficiencia en el procesamiento y con más eficacia para la detección de rostros en videos. La forma del detector propuesto consiste de un novedoso descriptor que codifica la información espacial y eventual para representar los patrones faciales en un grupo de cuadros seguidos, clasificadores boosting para elegir y aprender de manera automática las características más excluidas y una cascada de dichos clasificadores boosting para mejorar la velocidad en la fase de la detección. Los experimentos hechos en videos de data set internacionales enseñaron que el detector propuesto alcanza una eficacia mayor a la metodología de Viola y Jones, conservando sus grados de eficiencia, con menos números de operaciones. Específicamente, se aumentó la eficacia en un 10 % en comparación con el uso de características Haar y un 5 % a lo que se refiere al utilizar descriptores LBP. Lo nuevo de esta metodología está avalado por cuatro artículos publicados. Se trabajó una nueva metodología de seguimiento que se basa en la agrupación de detectores espacio temporales el cual, diferencia de otros enfoques se tiene que procesar las detecciones que se obtuvieron, considerada la salida del detector

espacio temporal que se proponen como partes de trayectorias iniciales seguras. Para unir las partes y formar las trayectorias finales de cada rostro, que sean resistentes en cuanto a movimiento y aspecto, se soluciona un problema de asociación de datos. Para adecuar el movimiento propusimos un filtro de Kalman que se adapta y se ajusta de manera dinámica a sus parámetros sobre la base de la seguridad del detector. El aspecto de cada fragmento se adecua usando el histograma de la representación espacio temporal computada para esa detección. En los experimentos realizados se evidenció que, mediante la asociación basada en detecciones espaciotemporales, se obtienen caminos de rostros más largos y un menor número de falsas trayectorias que usando detectores frame a frame. Lo nuevo de esta investigación científica de este método queda avalado en una publicación. Se diseñó una nueva representación de secuencias de rostros, que se basa en codificar el vector de Fisher de rasgos binarios, con el propósito de reducir el presupuesto computacional de la representación. Principalmente, se usa el descriptor BRIEF como rasgo binario, debido a su notoria simplicidad y almacenamiento sólido en memoria. Cuando los rasgos BRIEF se extraen de una forma densa y a múltiples escalas de todos los cuadros del video, se usa un método de análisis de componentes esenciales logístico para mostrar dichos descriptores binarios a un espacio vectorial de valores reales así poder utilizar la formulación clásica vector de Fisher basada en GMM, se reduce la dimensión de los descriptores locales e incluye la información momentánea agregando sus ejes espaciales al vector de rasgos proyectado. Los experimentos llevados a cabo en data set internacionales mostraron que la representación propuesta disminuye los tiempos de aplicación con una eficacia similar a la metodología relacionada en la literatura para la identificación/verificación de rostros. Particularmente, se enseñó en la evaluación donde se experimenta que la digitación del vector de Fisher de partes binarios BRIEF logra un apresuramiento de alrededor de 3 veces con respecto al tiempo de cómputo de la codificación vector de Fisher de rasgos SIFT, conservando una eficacia igual. La novedad científica de esta representación se valida por tres artículos publicados. Se diseñaron nuevos protocolos de evaluación para el data set internacional YouTube Faces que reflejan condiciones iguales a las existentes en las aplicaciones prácticas. Principalmente, se propuso un nuevo protocolo para verificar que se utiliza completamente de la data set y proporciona

un gran número de comparaciones, lo que ayuda a la evaluación más real de los métodos, considerando bajos índices de falsos que se aceptaron. También, se implementaron protocolos de reconocimiento en grupos abiertos como cerrados y se tiene en cuenta distintos tamaños de galería, tales como comparaciones de videos con videos y con imágenes. La evaluación de distintas metodologías de la literatura bajo los protocolos propuestos enseña una observación distinta cuando se utiliza el protocolo establecido. Las mejores respuestas obtenidas muestran que la eficacia de reconocimiento le falta bastante. Por eso, el desempeño de la metodología evaluada presenta una base para comparar próximas investigaciones en el reconocimiento de rostros en video. El aporte científico de los protocolos diseñados queda aprobado en una publicación. Para la validación de la ejecución de las consecuencias de esta investigación en escenarios reales, se realizaron pruebas en videos que vienen de las cámaras de video protección de Cuba, que fue parte de los resultados de un trabajo de tesis de doctorado en Ciencias Técnicas. En este trabajo se trabajó una plataforma experimental que permite realizar las pruebas en videos reales. Por otro lado, el aporte teórico de este trabajo está esencialmente en el desarrollo de un conjunto de metodologías que contribuyen a ejecutar el reconocimiento de rostros en videos de manera eficiente y con una eficacia al nivel de los métodos existentes. Estos métodos no solo tienen novedad científica en el país, sino que también son contribuciones en esta área del conocimiento a nivel internacional, como lo aprueban las publicaciones en revistas y congresos de alto prestigio en este tema. Estos nuevos métodos son la base teórica del conocimiento para solucionar el problema del reconocimiento de rostros en diferentes escenarios de video protección, lo cual es una necesidad del Ministerio del Interior (MININT). En este sentido, lo que significa la prueba de este trabajo se da en primer lugar, por la posibilidad de utilizar las metodologías propuestas en el desarrollo de sistemas propios para la identificación de rostros en aplicaciones de vídeo protección. Precisamente, es importante este tipo de aplicaciones para la defensa y el orden interior de cualquier país, la mayoría de los sistemas de video protección no se encuentran libremente disponibles en el mercado. Además, los sistemas de reconocimiento de rostros que existen en el mercado internacional son costosos, sin mencionar que en algunos casos existen restricciones para su adquisición y despliegue en Cuba. Especialmente, esta

investigación básica está dada a buscar resolver los problemas vinculados con la video protección, donde la necesidad es determinar o verificar de manera automática la identificación de las personas a partir de su cara, según el lugar donde se realice la vigilancia. Se prevé que las soluciones propuestas en esta investigación sean incrustadas en el sistema Xyma Safe Vision que trabaja la entidad con datos para la video protección, con lo que se le añade el valor agregado de identificación caras en los lugares que se analizan, y agregaría las herramientas con las que tiene el MININT, trabajadas por la entidad. Otras probables aplicaciones de interés de los resultados que se obtuvieron están grabadas en la identificación de caras en sistemas móviles o ingresados, donde necesitan algoritmos con bajas demandas digitales por las características de este tipo de dispositivos.

Talavera [13], presentó su trabajo denominado “Modelo para la detección de rostros en secuencias de video con baja calidad” en Arequipa, Perú. La utilización de las ciencias aplicadas actualmente está en varias labores. Cada vez se usan con mayor frecuencia una computadora, dispositivos móviles, cámaras de video, etc, con el fin de optimizar facetas, observar círculos populares y también verificar labores de delincuentes. Cada vez se ingresa mayor capacidad de videos y fotografías que son usadas frecuentemente en el área de visiones computacionales con la finalidad de trabajar temas de localización o análisis de labores humanas de manera autónoma. En estos temas, la que mayor realce tiene últimamente es cuando se interpreta una de secuencias de video usadas en vigilancia. Los dispositivos de vídeo son usados con efectos de monitorización están teniendo un mayor interés en grupos, entidades robustas, diminutas, negocios, etc. que están buscando conservar sus acciones, sustracciones y actividades ilegales por parte del personal que trabaja. Últimamente, el uso de las cámaras de vídeo es usadas frecuentemente para prever o avisar sobre acciones delincuenciales, inconvenientes de tránsito de vehículos, accidente, emergencia o temas de riesgo que pueden estar en el lugar. Por eso el aumento de la necesidad de agrandar los espacios de vigilancia para superar la protección, se observa las labores de más humanos con el fin de reconocer actividades extrañas y aumentar la demanda en lo que se refiere a escala y capacidad de las cámaras de vigilancia. La necesidad

lleva también al trabajo de la persona dedicada al manejo de las cámaras y/o estudio de las escenas, transformando esta labor en una fase que necesita bastante importancia por tiempos alargados. Sin embargo, esta fase que necesita de bastante importancia no puede ser llevada con mucho rigor por el trabajador, ya que se observa seguido en las secuencias de vídeo da cansancio visual del trabajador. Por ese motivo, los frame de las cámaras se controlan muy poco o nada. En algunos casos las grabaciones solo están como un archivo por si ocurre un evento delictivo o accidente. Las cámaras son herramientas con más valor si en vez de capturar de manera tranquila figuras, se utilizan para localizar hechos que necesiten ser atendidas a medida que pasan, y se tomen actos de prevención o corrección en tiempo verdadero. En ese sentido se necesita obtener aplicativos de vídeo vigilancia con fin de solucionar de manera automática el estudio del accionar de la persona tratando de ver futuros actos o identificar objetos para analizarlos. Una ventaja, es que los videos poseen información muy grande que una insignificante figura, estando en sistemas con procesamiento más grande y estático que se alcanza con la unificación de información obtenida de las diversas figuras o Frame en las tareas que son realizadas con el análisis de los frame, se resalta la detección de humanos, fase por la que se hace un estudio de las peculiaridades de los rasgos faciales de los humanos. Aunque todas las aplicaciones de la actualidad de reconocimiento de rostros han logrado un porcentaje en el grado de madurez, el trabajo y su ejecución en frames sigue teniendo un límite ya que no se puede controlar las condiciones en las que se trabaja. Teniendo como ejemplo, la fase de localización de caras en figuras obtenidas de los frame de un lugar exterior, tiene fases de luz con variaciones en la intensidad, la postura de los objetos no es igual, hay la existencia de obstrucción por parte de otro objeto, el ángulo de observación o la pequeña resolución de las imágenes que se obtienen son difíciles para procesar los algoritmos. Todas estas peculiaridades de un Fotograma o Frame hacen complicada la ejecución de las tácticas de localizaciones y reconocimientos de caras, las cuales se han diseñado de manera original para imágenes en entornos semi controlados, ya que la utilización de los dispositivos de vídeo vigilancia está en incremento, y se está desafiando a laborar con frame en tiempo reales es complicada para los aplicativos de localizaciones y reconocimientos de caras propuestas en la actualidad, en ese sentido, se plasma un novedoso esquema para

procesar la localización de caras que labora en frame con poca calidad (centrándose en la localización de caras que tienen poca calidad de imagen).

Salcedo [14] , en su investigación “Detección y Reconocimiento de Rostros en Imágenes Infrarrojas sobre Hardware Digital Dedicada” Nos habla sobre el uso de cámaras de vídeo para vigilancia, en donde la mayoría de dispositivos no poseen una buena calidad de captura estándar o similar, esto depende que no se cuenta con el recurso suficiente para poder adquirir cámaras de video con un resolución de imagen bastante elevada, en donde aquellas cámaras se modifican en el apresuramiento de captación, la resolución, la repercusión del acercamiento para capturas cerradas, capturas nocturnas o en ambientes oscuros, entre otros. Todas las características se modifican de según la índole de cada dispositivo, los cuales tienen influencia en los sistemas autónomos de vigilancia, así como el de la localización del rostro, ya que si posee una mala resolución no permite tener una verdadera detección de las características que pertenecen a un rostro, o no logran localizar características físicas propias de cada rostro, esto genera muchas veces errados positivos o errados negativos que complican la localización del rostro. Hay factores que modifican el modo de percepción del rostro, como la posición de la cámara que pocas veces es directo hacia el rostro, la luminosidad también es un factor que cambian la recepción del rostro ya que genera contornos que ocultan porcentajes de los rasgos faciales, otro problema son los artículos que llevan en el rostro, ya sea gorros, gafas, mascarillas, entre otros, que utilizan los delincuentes para no ser identificados. Varias de las tácticas que se han propuesto en la localización de rostros por varios autores, laburan exclusivamente en encontrar particularidades que alcancen a reconocer un rostro a base de algunos parámetros, es por ello que no logran obtener resultados favorables, esto se debe a la baja resolución del vídeo, la segmentación del objeto resulta ser complicado, esto se debe a que se tiene poca luz en los contornos y se pierde bastante información importante de las particularidades físicas de los objetos ya que no logran ser localizadas por las cámaras de video, la intensidad de luz es otro punto que cambia la visión en la perspectiva de los rasgos faciales así como el ángulo de la cámara de vídeo, todos estas particularidades son los inconvenientes fundamentales que aparecen en los sistemas de Detección de Rostros trabajados en secuencias de

vídeo.

Bernal [15] en su investigación denominada “Análisis de Métodos de Reconocimiento Facial Bajo el Sistema Operativo Android” nos comenta que, Actualmente las Tecnologías de la Información (TI) se trabajan livianamente de tal forma que varias de las ciencias aplicadas priman un rol fundamental en las entidades. Las ciencias de biometría ofrecen aguante a los procedimientos de las empresas desde los grados operativos hasta los niveles más altos, ayudando conjuntamente a tomar decisiones y a la protección de la información. Las entidades desde la más pequeña hasta la más grande desean proteger el paso a la información por ese motivo se desarrollan diversas técnicas biométricas para resguardar que se filtre la información. Una de las técnicas más investigadas actualmente es el reconocimiento facial, que se trata de la interconexión del hardware y software, de tal manera que optimiza el procedimiento de verificación del cliente, de diferente manera de otras tácticas biométricas, es muy cara la suplantación de datos. Para conseguir el reconocimiento de rostros se procesa la imagen, que es una forma más competente de seducir las metas y estructuras del entorno, dando acceso a la interconexión de humano-dispositivo. De este modo han salido distintas metodologías de reconocimiento de caras como es el ejemplo de la textura de la cara, por distancia de los puntos primordiales del rostro, por edificación de caras en 3D, entre otros; estas metodologías requieren la plataforma tecnológica para implementar, así como dispositivos con el espacio necesario. De tal forma que hay dispositivos emergentes con poco espacio, así como los equipos móviles, por lo que algunas metodologías de reconocimiento facial gastan bastante recurso y sale necesario el analizar qué metodología gasta menos recurso, con una tasa significativa aprobación.

Miranda, Gallego y Margolles [16] en su investigación denominada “Detección de Rostros Empleando MATLAB” nos comenta que, El reconocimiento de caras por medios computacionales, forma parte del área denominada biometría. La biometría proviene del vocabulario griego “bios” de vida y “metron” de medida. Es una aplicación de reconocimiento humano que se basa en particularidades físicas

(huellas dactilares, iris, forma de la mano, cara) y de actos (voces, firmas, dinámicas de tecleos o manera de andar), estas aplicaciones conllevan un fin específico y es de autenticar las personas y así evadir estafas, hurtos de información, no permitir el paso a conexiones y computadoras y como armamentos contra el terrorismo; también podemos verificar la identidad de un malhechor, ya que corrobora datos propios en cada ser humano. “Todos los individuos tienen rasgos únicos que se diferencian de los demás. La forma del rostro, la geometría de rasgos del propio aspecto como las manos, nuestras vistas y quizás la que se conoce más, las huellas dactilares, son algunos aspectos que los distinguen de los demás individuos”. Este tema se viene estudiando anteriormente y se considera hoy en día como la metodología ideal de identificación. La identificación de una persona a través de la cara, lo hace una manera muy segura y confiable al momento de implementar un mecanismo en la identificación de humanos, incluso llegamos a identificar estados de ánimo por los que se puede ver situaciones anormales. Resulta de interés aclarar las diferencias que existen entre los términos: reconocimiento, verificación e identificación. Reconocimiento es un término general, y no envuelve fundamentalmente la verificación y reconocimiento. Todos los sistemas biométricos hacen reconocimientos para “volver a reconocer” a un individuo que anteriormente fue ingresado. Verificación es un trabajo por el que el sistema biométrico trata de afirmar la semejanza declarada de una persona a través de la comparación de una muestra con varias planillas que se obtuvieron con anterioridad. La Identificación es un trabajo mediante la cual el sistema biométrico trata de determinar la identificación de un humano. Los datos biométricos se comparan contra todos los datos que se alojan en el data set. La identificación de “serie cerrada” se expone si el individuo figura en la base de datos, la de “serie abierta” donde no existe garantía de que el individuo figure en el data set. El sistema determina si la persona existe o no en la base de datos.

Dicho esto, se define como objeto de estudio: la fase de detección de rostros de personas. Objetivo general de la Investigación: implementar la detección de rostros de personas basado en el algoritmo de Viola[1]Jones empleando MATLAB Barriga [17], en su tema “Aplicación Práctica de la visión Artificial para el Reconocimiento de Rostros en una Imagen, Utilizando Redes Neuronales y

Algoritmos de Reconocimiento de Objetos de la Biblioteca Opencv” nos habla sobre el análisis de las herramientas de la visión humana que es el fin de la Visión Artificial o también denominada Visión por Computadora. La visión por computador explora la creación de aplicaciones que pueden identificar un objeto específico en una figura. Para las personas este trabajo es normal y se realiza seguidamente sin preocupaciones por el cómo; sin embargo, cuando pasamos este trabajo a un computador, se encuentran varios obstáculos que solucionar. Para definir la visión por computadora somos capaces de elegir por la definición de Marr (Marr D. 1982) en donde se encuentran rasgos necesarios que conducen a comprender la visión por computadora como el estudio de las fases asociadas a la visión normal, y que tiene como objetivo comprender estas fases y generar computadoras con capacidades iguales. Una de las ramas con más proyección en la Visión Artificial es el que se vincula con el Reconocimiento de la Biometría de Rostros, el cual se concentra en el mismo localizador que utilizan las personas para diferenciar un individuo de otro, el cual es: su cara.

Por lo que, uno de los fines primordiales es entender el abundante sistema visual de las personas y cómo se exponen las caras para conseguir separar identidades con precisión. La causa más tediosa en la fase de identificación de rostros, está en los grandes cambios que hay entre las diferentes imágenes de la cara de un mismo ser humano. Estas desigualdades se dan por los rasgos faciales, la variación de la luz, distintos peinados, maquillaje y accesorios, gracias a todos estos implementos un individuo puede parecer distinto en dos imágenes. Este trabajo se basa en analizar las distintas metodologías y algoritmos que facilitan localizar una cara en una figura, y ejecutar un sistema que se entrene para hacer el reconocimiento de un individuo específico en de la imagen.

Dang & Sharma [18] en su investigación “Review and Comparison of Face Detection Algorithms”, nos habla que, El ritmo en el desarrollo de la tecnología se debe a que la las computadoras se están volviendo inteligentes. estos inteligentes las computadoras conducen a la interacción computadora-humano, que es una nueva era ayudando en la exploración de varios campos. Detección de rostros, es decir el subdominio de la localización de objetos es uno de los temas que se basan en interacción ordenador-persona. La detección de objetos es una fase que trata

con la detección de instancias de objetos de una clase en peculiar (como personas, automóviles, edificios o caras) en una imagen o video.

Existen numerosas aplicaciones de detección de objetos. en esta próxima era como detección de rostros, detección de peatones etc. La detección de rostros es el primer paso básico para todos los métodos de análisis como el reconocimiento facial, la alineación facial, la cara modelada, verificación facial y seguimiento facial, etc. En primer lugar, la cara es detectada por la computadora entonces solo el más se puede hacer analisis. El procedimiento de detección de rostros es realizado en una imagen arbitraria que determina las caras en una imagen diferenciando los patrones formados por las caras los otros patrones y devuelve las dimensiones de cada cara.

Esto se puede lograr si el modelado facial y la segmentación facial son hecho con precisión. Esto también requiere considerar varios aspectos de apariencia facial como pose, resolución, enfoque, ruido, color, sombra y oclusión etc. El proceso general de reconocimiento facial se muestra en la Fig. 1. En las últimas décadas, la detección de rostros se está investigando mucho, ya que es algo trivial tarea para computadoras. Se proponen varios algoritmos para esto. tarea. En este artículo se discuten cuatro algoritmos básicos que se utilizan para la detección de objetos:

- 1) Características de Viola-Jones
- 2) 2.SMQT y SNOW Classifier
- 3) Detección de rostros basada en redes neuronales
- 4) Admite detección de rostros basada en máquinas vectoriales.

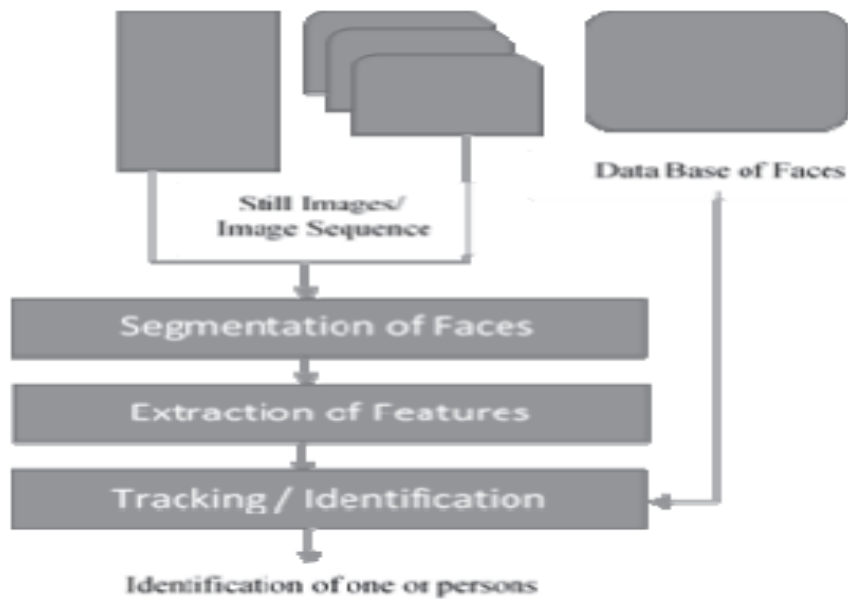


Figura 1: Proceso de identificación de rostros. Fuente: Dang & Sharma [14].

Después de la detección de rostros usando estos algoritmos, todos los los algoritmos se comparan sobre la base de la precisión y la recuperación Los valores de precisión y recuperación se calculan utilizando DetEval Software que toma dos conjuntos diferentes de archivos XML como aporte. El primer XML consiste en un rectángulo de verdad fundamental. dimensiones del valor del y el otro tiene las dimensiones de rectángulos detectados (cuadros delimitadores) de ubicaciones de rostros. Estos dos XML se comparan y las medidas de rendimiento se calculado. Tradicionalmente para la evaluación de la detección de objetos algoritmos, un solo archivo de detección y su terreno correspondiente archivo de verdad el recuerdo y la precisión, se puede calcular definido como sigue:

$$ROB = \frac{\text{N}^{\circ} \text{ de rectángulos detectados correctamente}}{\text{No. de rectángulos en la base de datos}}$$

$$POB = \frac{\text{N}^{\circ} \text{ de rectángulos detectados correctamente}}{\text{Número total de rectángulos detectados}}$$

La precisión y el recuerdo son las medidas utilizadas para medir la relevancia del algoritmo. Precisión también conocida como positivo valor predictivo que indica la

fracción de instancias recuperadas relevante y da información sobre las falsas alarmas. Recordar da la información sobre el número de objetos detectados o instancias relevantes recuperadas. Para el sistema perfecto cuanto más es cercano a 1 más precisos los resultados del algoritmo.

Justificación e importancia del estudio

La investigación es pertinente dentro del contexto ya que permite aumentar los conocimientos propios sobre las tecnologías actuales y con el fin de poder ayudar al avance tecnológico a nivel local, regional y nacional por ello se trabajó arduamente para para poder realizar todos los objetivos propuestos. Esta investigación también colabora con fines académicos ya que es necesario promover la cultura tecnológica en cuanto a biometría facial, dando un avance significativo de este tema, por otro lado, la inteligencia artificial que está en aumento y que sirve de herramienta en diferentes ámbitos sea en negocios, ámbito educativo, científico-experimental, médico y más.

Es relevante ya que en la actualidad se sabe que la inseguridad en el país es un problema que nos afecta a todos, por eso la importancia que tiene esta investigación se hace muy fuerte ya que se trabajó con el fin de que en cada parte donde se utilice esta investigación sea de bastante ayuda con la detección del rostro, dando una mayor aceptación en lo que es la detección ya que se tuvo que lidiar con bastantes inconvenientes para poder brindar un trabajo de calidad que será de beneficio y ayuda para muchísimas personas. Su acoplamiento y comprobación durante las fases de pruebas y operativas se considera necesaria y fundamental para poder avanzar con las etapas del reconocimiento facial además de salvaguardar los ecosistemas gracias a la capacidad de disponer de una herramienta que no contamina el medio ambiente.

Dado el alcance que tuvo esta investigación se acota que está presente investigación se trabajó con personas que conocen bastante del tema en donde se pudo destacar el compromiso y la profesionalidad que tienen y para ello se tuvieron que fijar fechas exactas en donde se presentaron avances de la investigación la cual se llegó a una instancia donde se pudo decir que el trabajo realizado fue de manera responsable y aceptable.

Es viable la realización de este proyecto ya que va a reducir costos porque no se tendría que comprar el sistema al extranjero en donde se haría más costoso y más tedioso al momento de conseguirlo. Del mismo modo, se considera también básico para poder realizar tareas de análisis y gestiones de la calidad de la detección de rostros.

Limitaciones de la investigación

En el proceso de elaboración de la investigación se presentaron distintas limitaciones las cuales hicieron que la investigación sea más compleja y de mayor demanda de la que se esperaba, para esto se enumeró las limitaciones que se obtuvieron durante el proceso de la investigación:

- 1) La baja calidad de las cámaras las cuales no permitieron tener una imagen clara del rostro de las personas.
- 2) El ambiente en donde se localiza la cámara ya que puede estar muy oscuro o claro.
- 3) La posición de las personas, ya que mientras más sea la rotación del rostro, la red neuronal va a reducir su rendimiento.
- 4) Los accesorios que llevan las personas, ya que al llevar accesorios en el rostro dificulta la detección del rostro
- 5) Tiempo limitado para el proceso de la investigación

1.2. Formulación del problema

Después de analizar la situación problemática actual se realizó el siguiente problema en manera de interrogación: ¿Cómo reconocer rostros a través de las cámaras de seguridad interna?

1.3. Hipótesis

Con las técnicas de visión artificial y procesamiento de imágenes desarrolladas en el lenguaje de programación Python, se podrá reconocer rostros en las cámaras de seguridad interna.

1.4. Objetivos

Objetivo general

Desarrollar un sistema de reconocimiento de rostros aplicando técnicas de inteligencia artificial y procesamiento de imágenes en cámaras de seguridad interna.

Objetivos específicos

- Diseñar un software utilizando las herramientas de python y la librería opencv que me permita enlazar el software con la cámara.
- Diseñar una red neuronal que permita reconocer rostros, utilizando las técnicas de visión artificial y procesamiento de imágenes.
- Diseñar un software que permita almacenar los rostros detectados.
- Evaluar resultados.

1.5. Teorías relacionadas al tema

¿Qué es la detección facial?

La detección facial es analizar distintas fotografías y videos e identificar dónde existe un rostro, sin importar a quién pertenece el rostro. Hay varias técnicas que permiten la detección de caras, como la longitud entre los ojos, distancia entre nariz y boca. Se han trabajado distintos algoritmos que, en porcentajes elevados, reconocen los rostros en una fotografía o video y a través de esto se realiza el reconocimiento facial diferentes sistemas de seguridad. En la actualidad la detección facial es parte del reconocimiento facial en diferentes equipos móviles y sistemas de seguridad. El proceso de reconocimiento facial no siempre consigue un buen resultado porque depende de diferentes factores físicos y de la obtención de la imagen tales como: la variación en la luz que recibe el rostro del individuo, el ángulo en que la fotografía fue tomada y las distintas resoluciones de la cámara empleada [19].

Haar Cascade

Existen varias técnicas que nos facilita aprender el reconocimiento de un objeto en una figura o fotograma de vídeo. La técnica cascada es un método de clasificación definida por Viola Jones, la cual se centra en la agrupación de varios clasificadores frágiles, cada uno analiza una parte opuesta de una figura o fotograma. Se consideran débiles porque tienen bastante probabilidad de encontrar una cantidad considerable de falsos positivos, pero al combinarse los resultados, son muy elevados. El objeto de la clasificación proyecta la figura según el costo de las figuras, que se manifiesta al conjunto de píxeles que se encuentran en una ventana de localización cerrada. Por tal motivo, se dice que hay muchos estímulos para usar figuras lugar de píxeles claramente en acción de la intensidad. El algoritmo usa ventanas con dos, tres y cuatro rectángulos con el mismo tamaño y dimensión. En cada ventana ejecuta la función clasificadora, que se estima como la suma de todos los píxeles que están dentro de los rectángulos blancos y se quita con la suma de todos los píxeles del rectángulo sombreado.

De acuerdo con la figura 2, se señala los rectángulos de igual dimensión donde se aplica el método Haar Cascade:

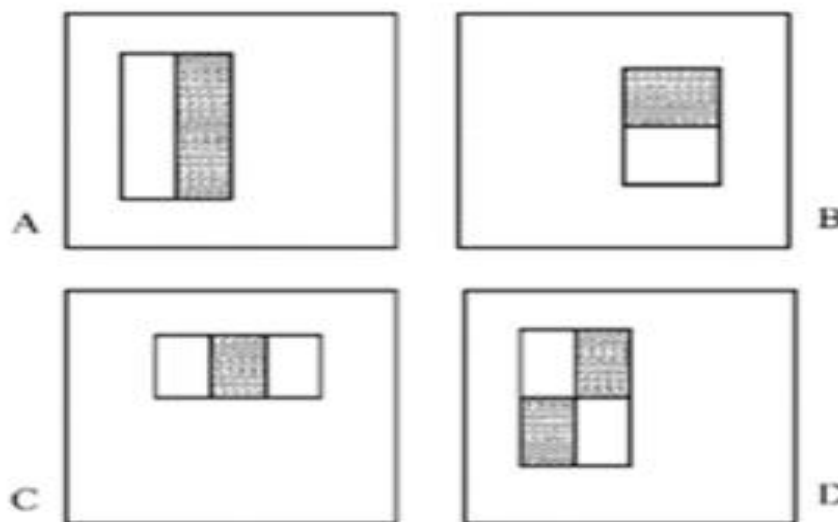


Figura 2: Feature de ventanas del Método Haar Cascade. [20].

Aquí encontramos características del objeto analizado, donde se compara la imagen positiva con la imagen negativa y hallar igualdades. Para entrenar las imágenes, necesitamos un data set con imágenes positivas, donde estará el objeto que desea detectar, además un data set de imágenes falsas donde no está el objeto

a detectar. Aquí se realizan todos los pasos en mención y se escogen las figuras más sustanciales. Esto requeriría una gran capacidad de memoria y por tal motivo se sugiere manejar Adaboost, ya que su función primordial es hallar fragmentos de las figuras que son relevantes del objeto. Cuando se aplica todas las features relevantes al mismo tiempo al momento de reconocer un objeto, sería malo en tiempo y presupuesto, por lo que se aplica el método de haar cascade. Esto se debe a la función recursiva de la figura integral que ayuda con un veloz análisis de las figuras donde podríamos hallar el objeto. En el hipotético hecho de que fallara una subregión de la imagen, se descarta la ventana. Caso opuesto, si aceptara, pasar a una siguiente fase. Cada fase tiene un orden de prioridad; Al principio se encuentran las más importantes del objeto. Se considera que cuando menor sean las etapas se definen, será más gastos de recurso necesario para detectar si en una subregión se encuentra o no el objeto. [20].

CASCADE TRAINER GUI

La herramienta Cascade trainer GUI ayuda a entrenar, testear y mejorar los ejemplos de clasificación de cascada. Principalmente brinda una herramienta para obtener un data set de datos con figuras que pertenecen al objeto a detectar y otro data set con imágenes que no pertenecen con dicho objeto, mediante estas fuentes de datos obtenemos como acción un modelo de cascada en formato XML. Cuando se trabaja con estos mecanismos se debe ajustar algunos parámetros y valores. Lo cual genera diferentes modelos y compara los resultados. A continuación, detallamos los parámetros más resaltantes para configurar un entrenamiento de clasificación:

- Cantidad de pasos: Cantidad de veces que se repite el modelo mientras va entrenando. Se tiene que encontrar un número considerable de repeticiones a fin de eludir el overfitting (memorizar datos).
- Tamaños de buffers para valores e índices de memorias: Nos facilita señalar la cantidad de Megabytes que van a utilizar como buffer mediante los entrenamientos. Según el hardware que se dispone, para todos los entrenamientos ejecutados se configuró con una cantidad de 4096 Mb.

- Tamaño de las figuras: Esta herramienta sólo permite que los datos de entrenamiento tengan iguales dimensiones, si no fuera el caso, solo analiza la porción de la figura que le compete con los parámetros ajustados (en pixeles). Para ejecutar el entrenamiento, en base a los datos disponibles, se ajusta los valores con la siguiente cantidad: 24px de ancho y 24px de alto.
- Tipo de algoritmo utilizado: Se utiliza el algoritmo haar cascade detallado anteriormente. Se utilizó en varios entrenamientos, con diferentes sets de datos, donde se obtuvo diferentes modelos, que fueron usados como inicio de la aplicación de visión por computadoras y se obtuvieron ver algunas conclusiones sobre qué tan efectivo fue en las diferentes pruebas realizadas. [20]

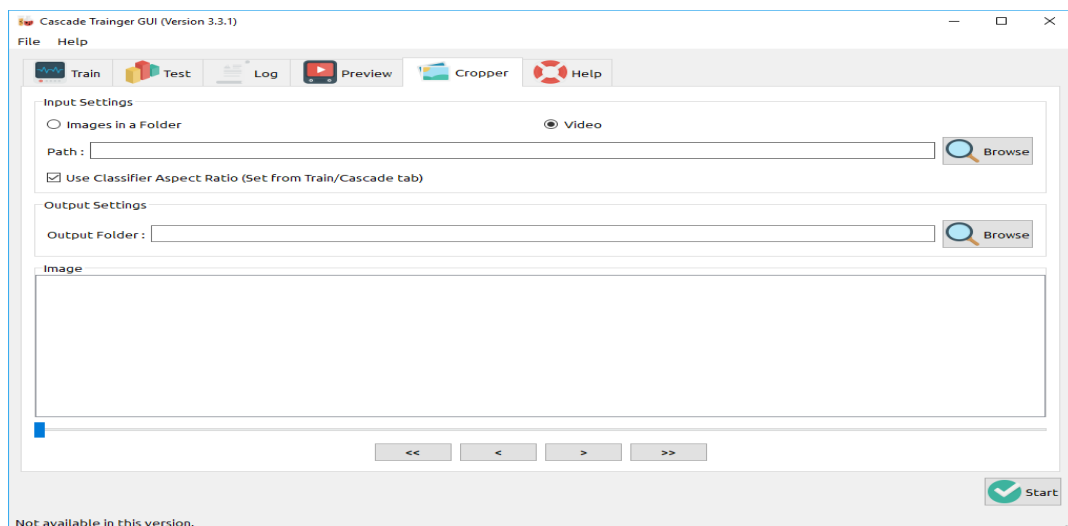


Figura 3: Entrenamiento de imágenes positivas y negativas. [20]

OPENCV EN PYTHON

Para aplicar el algoritmo de Haar Cascade es necesario usar librerías de visión artificial como OpenCV, a través de su ejecución en un sistema desarrollado en Python, lo cual ayudó a entrenar, testear y aumentar la calidad del ejemplo de clasificación usando Haar Cascades. Para mostrar en forma visible los objetos que han cumplido todos los procesos de selección, no se enfoca en la descripción de todas las funciones de esta librería, ni en el scrip que se implementó, sino que explica a groso modo los conceptos que se encuentran debajo de la lógica involucrada en el mismo. [20]

La detección de rostros en videos es muy importante para diversas aplicaciones de seguridad. Se necesita diferentes pasos que se ejecutan empezando por la entrada al sistema de una secuencia de imágenes o cuadros que tienen rostros de personas que se quiere detectar.

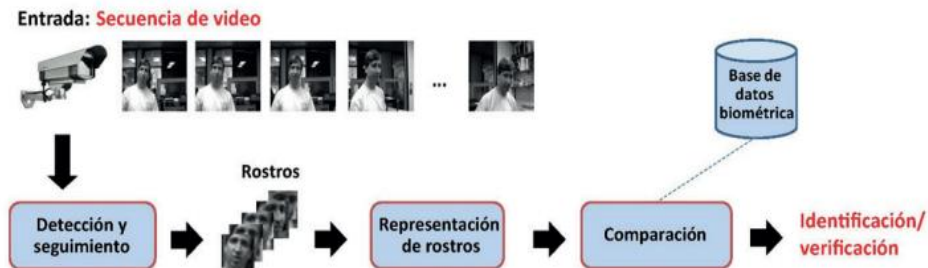


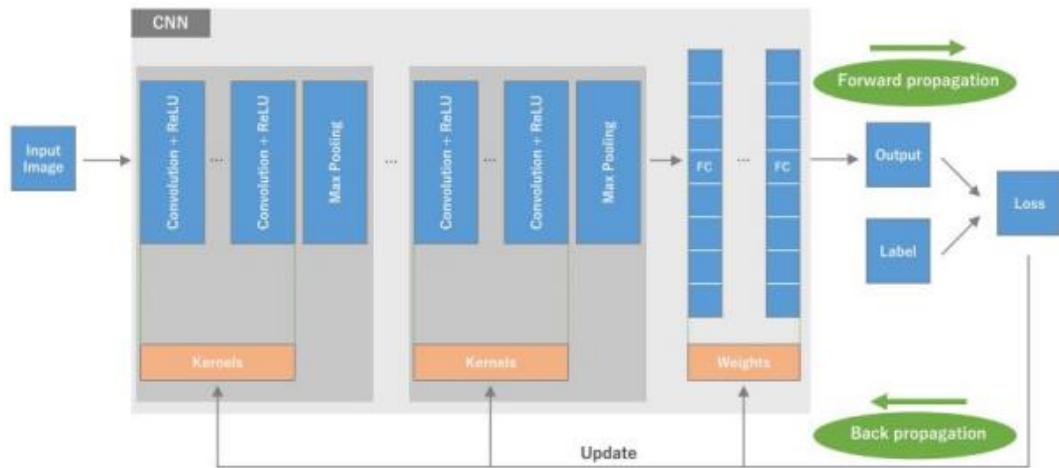
Figura 4: Esquema general de reconocimiento de rostros en videos 2019 [21].

En primer lugar, se determina, mediante la detección y el seguimiento, la ubicación del rostro en cada uno de los cuadros del video; después se pasa a la representación de las secuencias, y por último, mediante la comparación, se obtiene una respuesta de la identificación/verificación de los rostros analizados [21].

Red neuronal convolucional

La red neuronal convolucional (CNN) es una clase de red neuronal artificial que se ha vuelto dominante en varias actividades de visión por computadora. CNN está diseñada para entender de manera autónoma y se pueda adaptar a las jerarquías espaciales de las características a través de la propagación hacia atrás mediante el uso de múltiples bloques de construcción, como capas de convolución, capas de unión y capas totalmente unidad. La red neuronal convolucional es una elaboración matemática que normalmente contiene tres tipos de capas (o bloques de elaboración): convolución, unión y capas totalmente unidad. Estas dos primeras, capas de convolución y unión, extraen las características, mientras que la siguiente, una capa completamente conectada, mapea las características que se han extraído en la salida final, como la selección. Una capa de convolución genera un papel primordial en CNN, que se forma de diversas operaciones matemáticas, como la convolución, un modo especializado de operación lineal. En las figuras digitales, las cantidades de los píxeles se juntan en una cuadrícula bidimensional (2D), quiere decir, una matriz de números, y una diminuta cuadrícula de parámetros llamada

kernel, es el que extrae las características optimas, se aplica en cada pose de la imagen, lo que genera que las CNN sean muy buenas para procesar imágenes, ya que una característica se puede encontrar en cualquier parte de la imagen. A través



que una capa alimenta su salida a la siguiente, los grupos extraídos pueden volverse más complicadas de forma jerárquica y progresiva. La fase de optimizar parámetros como los kernels se llama entrenamiento, el cual se hace para reducir la diferenciación de las salidas y las etiquetas de verdad del terreno mediante un algoritmo de optimización llamado retro propagación y descenso de gradiente, entre otros.

Figura 5. Arquitectura de red neuronal convolucional (CNN) y el proceso de entrenamiento. [22]

La arquitectura de CNN tiene diferentes componentes básicos, como capas de convolución, capas de unión y capas totalmente unidas. Una arquitectura típica consiste en repetir una pila de diferentes capas de convolución y una capa de agrupación, seguidas de una o más capas totalmente unidas. El paso en el que los datos de entrada se transforman en salida a través de estas capas se llama propagación hacia adelante. Chavez Leyton [22]

Estructura de una Red Neuronal

Cuando se ha comprendido la teoría de una neurona artificial, necesariamente se conoce cómo estas se entrelazan entre ellas en un tejido de neuronas artificiales. Un tejido neuronal sintético está “compuesta por nodos o unidades que se conectan

por vías que son dirigidos”, siendo los nodos las neuronas artificiales y las vías son los medidores de vínculo entre sí. Estos vínculos son los encargados de difundir cómo se encuentra activada una neurona salida a otra neurona llegada, los vínculos poseen un peso vinculado que señala la vehemencia de conexiones entre las dos neuronas vinculadas. Desde la perspectiva computacional, una red neuronal es un ejemplo matemático que da marcha a varias funciones matemáticas sobre una información de origen, los cuales se procesan y transforman a través de su avance por las distintas capas de la red hasta originar una salida. La salida puede ser un número o un listado de cantidades. La forma más común de una red neuronal incluye:

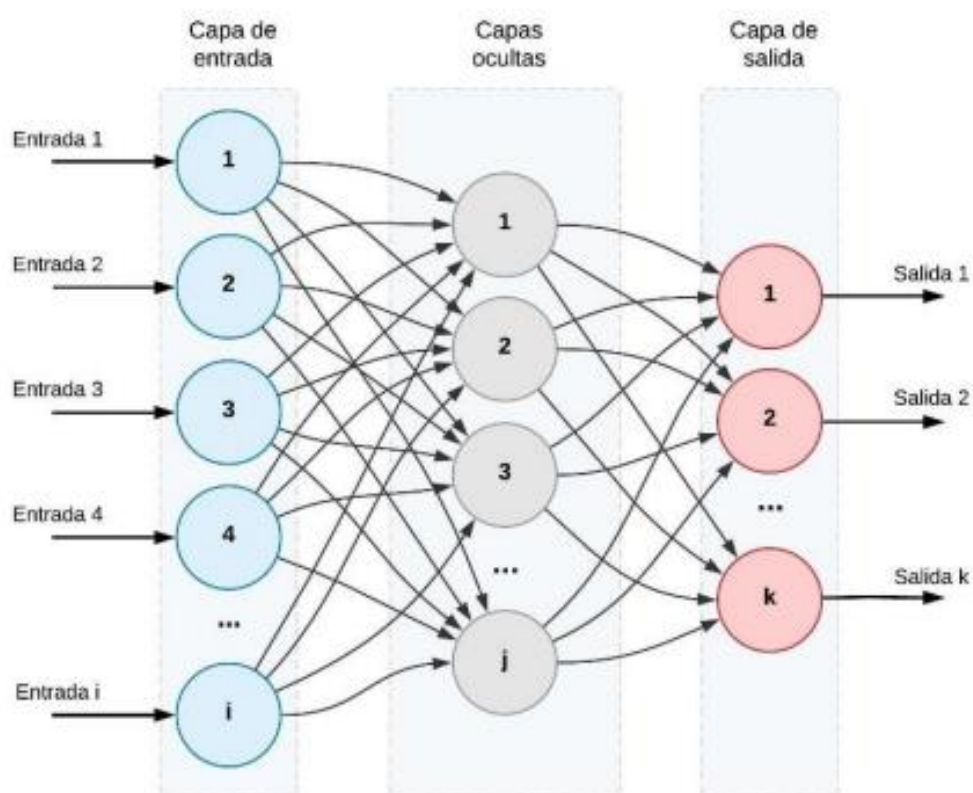


Figura 6. Red de neuronas [8]

Capa inicial: capa que acoge la información de origen. Hay una neurona para cada cualidad del modelo procesado que se quiere poner en marcha. En el tema del procesamiento de figuras, hay una neurona de origen por cada píxel de la figura de inicial (cuando ya se ha convertido los píxeles a cantidades numéricas).

- Capas ocultas: capas que están entre la capa de origen y la capa de destino. Esas capas se encargan de detectar patrones, que permitirán al modelo clasificar las entradas que se han recibido.
- Capa de salida: La última capa de la red, que brinda el dato o datos de salida que genera el tejido neuronal.

Hay una gran diversidad de tejidos neuronales artificiales que se clasifican a través de su arquitectura. El ejemplo más sencillo de un tejido neuronal es el perceptrón simple (se conoce en inglés como Single Layer Perceptrón), propuesto por Frank Rosenblatt en 1958, que se entiende como un tejido de una única capa de neuronas u, observado de otra manera, donde las entradas están vinculadas de manera directa a las salidas. El perceptrón es un instrumento que sirve para resolver inconvenientes donde el objetivo es seleccionar inconvenientes linealmente separables. Sin embargo, tiene bastantes límites si se plasma como respuesta a otros asuntos más complicados. Por tal motivo, los tejidos neuronales artificiales han resuelto hacia esquemas más complicados, ya sea por la cantidad de neuronas que tienen como por la naturalidad de éstas y sus Inter vínculos, como los tejidos neuronales convolucionales, las redes neuronales recurrentes, los autoencoders, los mapas autoorganizados de Kohonen, etc. [8]

Definición de Términos Básicos

Biometría: Se trata del análisis de procesos automáticos para la identificación única de personas que se basan en bastantes rasgos conductuales o rasgos físicos intrínsecos.

Método: Es la manera de ejecutar algo en forma sistemática, planificada y/o estructurada.

Metodología: Es un grupo de procesos racionales que se utilizan para lograr uno o varios objetivos que rige un estudio científico, una exposición doctrinal o labores que necesiten habilidades, conocimientos o cuidados esenciales. Frecuentemente se define como el análisis o selección de un método adecuado aplicable a determinado objeto.

Memoria de Video: Memoria semiconductora (RAM) donde se guarda la imagen

digitalizada. Normalmente la transformación analógica digital y la memoria de video se combinan en un módulo que se conoce como frame grabber(captura de imágenes).

Imagen. Es la delegación de un elemento o cosa que se puede observar de manera fácil. La cual la componen diferentes tonalidades donde nos brinda un ejemplo de hechos reales.

Pixel. Es el fragmento más pequeño de una figura digital. Cada pixel es un fragmento homogéneo, y con el grupo de pixeles con distintas variaciones de color brindan como resultado una imagen.

Recurso computacional. Es cualquier integrante físico, que se encuentra unido a un sistema de información con capacidad limitada. Es la parte interior o parte hardware donde se está trabajando un sistema.

Procesamiento de imágenes. Es una técnica para mejorar los aspectos de las figuras o fotografías, donde resalta la parte fundamental o mejorando la calidad de la figura digital.

Ambiente no controlado. Es una zona donde las condiciones no son estáticas y no permite manipular las intensidades de la luz.

II. MATERIALES Y MÉTODO

1.1. Tipo y Diseño de Investigación

Esta investigación es de tipo tecnológico – aplicada, porque se utilizarán las tecnologías necesarias para la realización de esta investigación, además de ser una investigación cuasi experimental porque se trata en elegir los grupos de la muestra en los que se prueba la variable sin un tipo de prueba aleatoria o proceso de preselección.

1.2. Variables, Operacionalización

En la siguiente tabla (**TABLA I**) vamos a describir las variables dependiente e independiente, incluyendo sus componentes para lograr la detección de rostros:

TABLA II
Operacionalización de la variable dependiente

Variable de estudio	Definición conceptual	Definición operacional	Dimensiones	Indicadores	Ítems	Instrumento	Valores finales	Tipo de variable	Escala de medición
Detección de rostros	Es analizar distintas fotografías y videos e identificar dónde existe un rostro, sin importar a quién pertenece el rostro	Representa la cantidad de rostros que se detectan dentro de una imagen o un frame	Precisión Revocación Valor F	Es el equilibrio de muestras correctamente diagnosticada del total de muestras evaluadas. Es la proporción de muestras correctamente diagnosticadas para cada tipo de carencia de calidad. Nos permite corregir el error de la distancia en los casos en los que la revocación (R) y la precisión (P) se compensan, brindando así una medida de precisión al algoritmo con el que se va a trabajar	$P = \frac{\#Taciertos\Omega\#100}{\#Tmuestras}$ $R = \frac{\#Taciertos\Omega\#100}{\#Taciertos}$ $F = 2 \frac{Precisión * Revocación}{Presición + Revocación}$	Técnica: Revisión Documental Instrumento: Ficha de registro de datos	Contabiliza cuantos segundos dura el Procesamiento de las muestras	numérica	Malo: 0% – 9% Regular: 40%-59% Bueno: 60% -79% Muy bueno: 80% -95% Excelente: 96% a mas

Nota: Se define la variable dependiente con las dimensiones e indicadores que se van a implementar. Fuente: Elaboración Propia

1.3. Población de estudio, muestra, muestreo y criterios de selección

Población: La población está compuesta por una base de datos de 100 videos con una duración promedio de 4 – 5 minutos. La base de datos de los videos es propia.

Muestra: La muestra se está calculando aplicando la muestra estadística finita en una base de datos

$$n = \frac{N * Z^2 a * p * q}{d^2 * (N - 1) Z^2 a * p * q}$$

Donde:

- N = Total de la Población
- $Z\alpha = 1.96$ al cuadrado
- p = proporción esperada (en este caso 0.05)
- q = 1 – p (en este caso 1-0.05 = 0.95)
- d = precisión (será un 5%).

$$n = \frac{100 \times 1.96^2 \times 0.05 \times 0.95}{0.05^2 \times (100 - 1) + 1.96^2 \times 0.05 \times 0.95}$$

En donde nuestra muestra es un total de 42 videos.

1.4. Técnicas e instrumentos de recolección de datos, validez y confiabilidad

Se hará uso de los siguientes instrumentos para la recolección de datos:

Observación: A través esta técnica se quiere analizar el accionar de la red neuronal aplicada en función de su rendimiento.

1.5. Procedimiento de análisis de datos

Observación:

En la observación se realizará diferentes pruebas para medir el comportamiento de los algoritmos aplicados en esta investigación para después tratar de ver las mejoras con el objetivo de tener buenos resultados en donde se requerirá la ayuda de un experto el cual nos hará ver las fallas en las pruebas a realizar.

Indicadores en variable independiente:

- **TIEMPO (vector de iniciación):** Contabilizar cuantos segundos dura el Procesamiento de las muestras.
- **TECNICA:** Calcula cuantos segundos le toma a cada técnica realizar el diagnostico
- **Indicadores en variable dependiente:**

Precisión: Es el equilibrio de las muestras correctamente diagnosticada del total de muestras evaluadas.

$$P = \frac{\#Taciertos\Omega\#Tmuestras}{\#Tmuestras}$$

Revocación: Es la proporción de muestras correctamente diagnosticadas para cada tipo de carencia de calidad

$$R = \frac{\#Taciertos\Omega\#Tmuestras}{\#Taciertos}$$

Valor - F: Nos brinda el acceso para emendar el error de la distancia en los casos en los que la revocación (R) y la precisión (P) se compensan, brindando así una medida de precisión a la red neuronal con el que se va a trabajar.

$$F = 2 \frac{\text{Precisión} * \text{Revocación}}{\text{Presición} + \text{Revocación}}$$

1.6. Criterios éticos

La presente investigación se realiza bajo unos criterios éticos que todo investigador debe tener en cuenta.

Entre los principios éticos que se va a desarrollar tenemos los siguientes:

Confidencialidad: Los códigos de ética hacen énfasis ante la seguridad y el resguardo de los participantes como la identidad, como también los informantes de la investigación.

Derechos de Autor: Todo documento e información utilizada para el desarrollo de tema de tesis estará referenciado y citados con sus respectivos autores como participante del trabajo

Criterios de Rigor Científico

Originalidad.

Se detallará de donde provienen las fuentes bibliográficas correctamente analizadas y consultadas, con el fin de ser ponderados en fundamento al porcentaje resultante del programa anti-plagio.

Veracidad.

La documentación expuesta en el presente trabajo, se exhibirá sin manipulación alguna. Puesto que, será totalmente veraz y confiable.

III. RESULTADOS Y DISCUSIÓN

2.1. Resultados

VARIABLE INDEPENDIENTE

Resultados en Tablas y Figuras.

La detección de rostros mediante cámaras de interiores se realizó con la librería de OpenCv en la plataforma de Python en este capítulo se muestran las tablas y figuras con las pruebas realizadas en la interfaz final y se describen los resultados que se han obtenido en dichas pruebas, en donde se han medido el porcentaje de la precisión, revocación y el valor final.

Los videos fueron conseguidos de cámaras de seguridad en ambientes semi controlados con variación en diferentes aspectos, ya sea de iluminación, pose y características de diferentes personas.

Teniendo en cuenta que el data set contenía 42 videos con imágenes de diferentes personas por video.

Precisión:

Esta prueba se realizó en diferentes ambientes, durante los diferentes horarios del día, contando con la red entrenada, estas pruebas se realizaron en diferentes horarios que se mencionaron anteriormente, estas pruebas se realizaron en ambientes semi controlados.

Para la prueba de detección de rostros se utilizaron videos almacenados en nuestro data set donde figuran distintas personas para los diferentes horarios, la cual son necesarias para poder obtener un porcentaje elevado de detección de rostros, para poder hallar este porcentaje de precisión, se trabajó la siguiente formula:

$$P = \frac{\#Taciertos\Omega\#100}{\#Tmuestras}$$

TABLA III

Escala Valorativa de cuantificación de indicadores (Variable Independiente).

CRITERIOS	MAÑANA	TARDE	NOCHE
Tamaño de la muestra	110	100	100
Nº de Aciertos	106	97	95
Nº imágenes no detectadas	4	3	5
Porcentaje	96.3	97	95

Nota: Definimos los diferentes criterios utilizados para la precisión, teniendo en cuenta los distintos escenarios, con los resultados obtenidos. Elaboración Propia

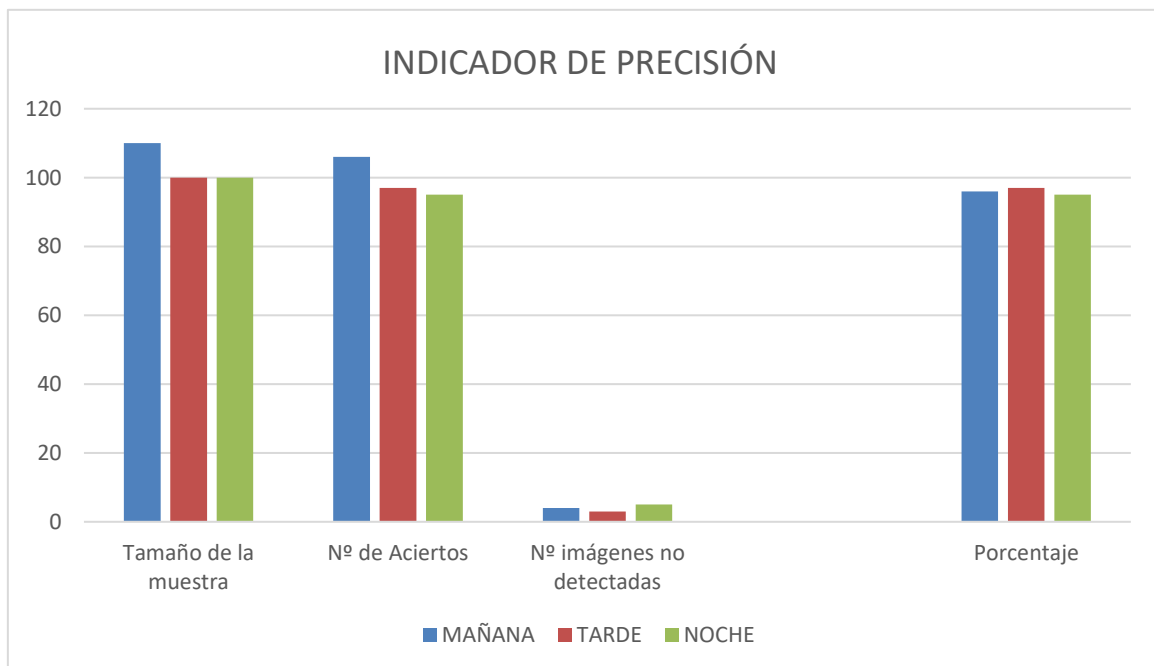


Figura 6: INDICADOR - PRECISIÒN. Elaboración Propia.

Interpretación:

Como se puede observar en las tablas y gráficos mostrados, se hizo una

comparación de los diferentes criterios de evaluación en los diferentes ambientes, en donde nos podemos dar cuenta que se ha obtenido un porcentaje bastante elevado con respecto a la precisión sobre la detección de rostros.

- **Revocación:**

Para la realización de este criterio se contó con imágenes extraídas desde nuestra data set con videos en diferentes ambientes, durante los diferentes horarios del día, contando con la red neuronal ya entrenada, estas pruebas se realizaron en diferentes horarios que se mencionaron anteriormente, estas pruebas se realizaron en ambientes semi controlados.

Para la realización de este criterio tenemos en cuenta el número total de aciertos y muestras, para obtener las muestras correctamente diagnosticadas.

Para la prueba de detección de rostros se utilizaron videos almacenados en nuestro data set donde figuran distintas personas para los diferentes horarios, la cual son necesarias para poder obtener un porcentaje elevado de detección de rostros, para poder hallar este porcentaje de revocación, se trabajó la siguiente formula:

$$R = \frac{\#Taciertos\Omega\#100}{\#Tmuestra}$$

Tabla IV
Pruebas de Revocación.

CRITERIOS	MAÑANA	TARDE	NOCHE
Tamaño de la muestra	110	100	100
Nº de Aciertos	106	97	95
Nº imágenes no detectadas	4	3	5
Porcentaje	96.3	97	95

Nota: Resultados de las pruebas de revocación en los diferentes ambientes en los que se han realizado las pruebas. Elaboración propia

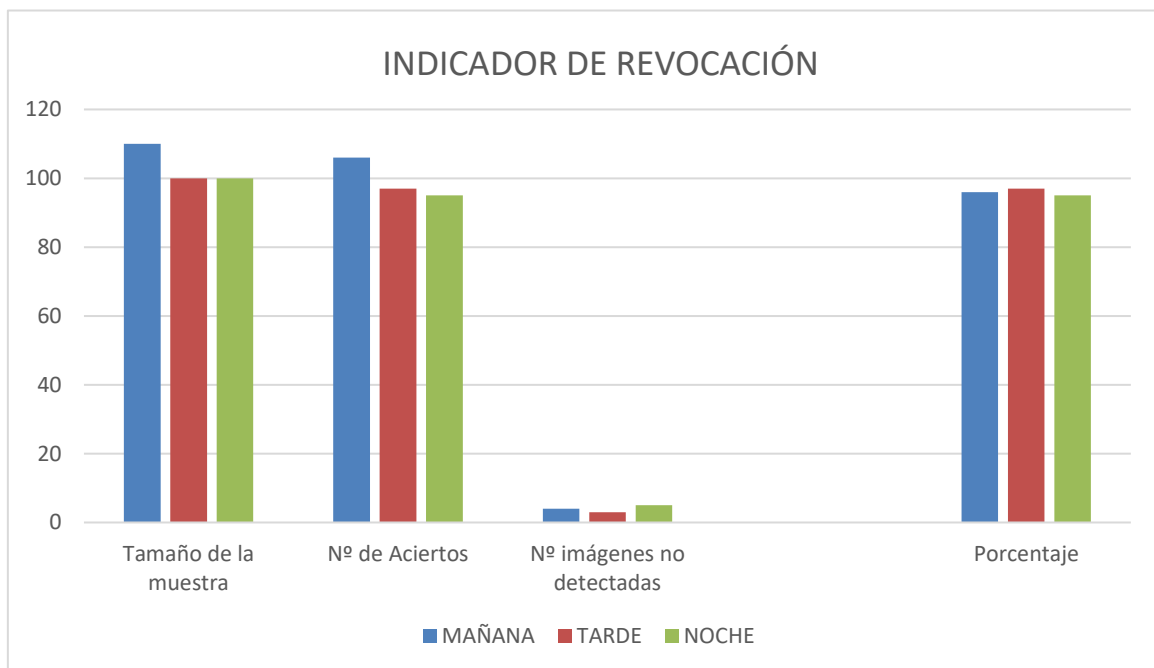


Figura 7: Gráfico de Revocación. Elaboración Propia.

Interpretación:

De la misma manera se obtuvieron resultados muy elevados en lo que corresponde al porcentaje de revocación, dando a entender que el algoritmo y la técnica utilizada han mostrado resultados favorables para el proceso de investigación.

VALOR F:

Para este criterio se realizó el porcentaje entre la precisión y la revocación para poder obtener el valor final, de esta manera se puede obtener un análisis mas exhausto sobre la eficacia que tiene el método estudiado , para este criterio utilizamos la siguiente formula:

$$F = \frac{2 * Presición * Revocacion}{Presición + Revocación}$$

Tabla V.

Resultado Valor F. Fuente: Elaboración propia.

CRITERIOS	PRUEBA1	PRUEBA 2
Precisión	96.3	97
Revocación	100	100
Valor F	98.1	98.4

Nota: Resultados de las pruebas de revocación y precisión que se han realizado las pruebas, del cual se obtiene su porcentaje final

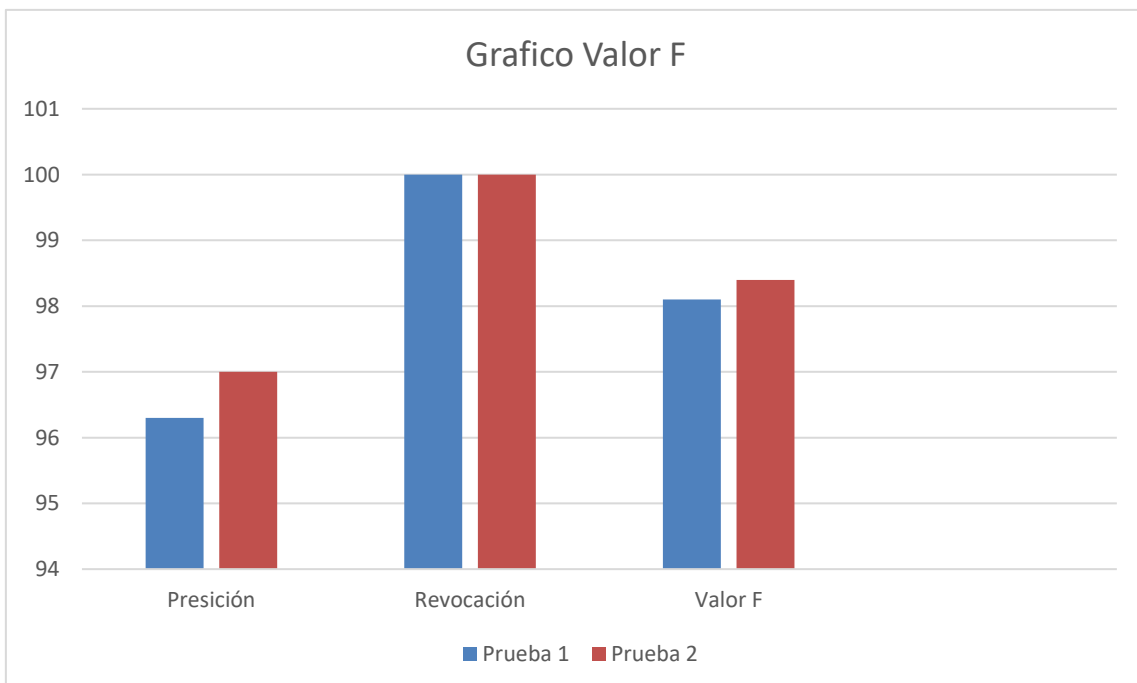


Figura 8: Gráfico de Revocación. Fuente: Elaboración Propia.

Interpretación:

Como se puede observar en la tabla y gráfico del Valor F, podemos observar la comparación de resultados entre la revocación y la precisión, en donde estamos obteniendo el porcentaje final, al apreciar los resultados en los valores finales podemos analizar que se han obtenido resultados muy favorable para la investigación, esto nos quiere decir que el entrenamiento de la red neuronal en diferentes ambientes ha resultado exitoso, lo cual nos brinda un avance para futuras investigaciones sobre este tema.

2.2. Discusión

Los resultados obtenidos en el criterio de precisión y revocación, nos indican que la red neuronal entrenada tiene una cierta deficiencia cuando se trabaja en un ambiente con poca luminosidad, tal es el caso de los videos obtenidos el horario nocturno por la poca luz que se obtiene, dichos resultados nos han servido para poder analizar las ventajas y desventajas que tiene nuestra red neuronal y así poder obtener mejoras para trabajos futuros.

Para comprender sobre la aplicación de la prueba, se tiene que entender que entradas se utilizan. Para esta prueba, las entradas fueron, una serie de videos de un diferentes ambientes y lugares que hoy en día utilizan para obtener seguridad. Además, se requirió de la salida un resultado que se obtuvo a partir de la red neuronal convulsional en base a los datos del entrenamiento sobre cada rostro que se desea detectar en el vídeo, por tal motivo, se trabajó con un modelo para la detección de rostros, estos modelos se entrenaron con distintos data sets de imágenes positivas y negativas. A través de la función CascadeClassifier del módulo cv2 de la librería OpenCV, en el código del sistema. Después se procesa el flujo de datos de vídeo y por cada frame detectado se hace lo siguiente:

Se transforma la imagen capturada a escala de grises, ya que los datos fueron entrenados con data sets de estas características. Esto se realiza principalmente para disminuir el procesamiento y hacer más fácil la detección en tiempos relativamente acordes.

Se consigue las coordenadas de distintas secciones donde el clasificador explicado previamente en base al modelo detecta el objeto con características

similares al rostro que se desea obtener. Esta función es importante en el avance, y se identifica como detectMultiScale en el módulo de la librería, ya que los rostros que detecta podrían tener distintos tamaños. Esto se realiza para cada clasificador del cual deseamos detectar rostros.

Cuando se obtienen todas las coordenadas de los rostros localizamos en el frame, procedemos a dibujarlos en el frame original (a color), se renderiza un rectángulo en la sección correspondiente.

Por último, se muestra el frame generado, con los rectángulos que identifican los rostros en una ventana para su visualización. A continuación, se puede ver un ejemplo del resultado.



Figura 9: Primera Prueba de detección de rostros. Elaboración propia

En esta primera prueba de la detección de rostros podemos observar que para la red neuronal se la hace complejo detectar rostros cuando se obtiene bastante luminosidad, la pose de la persona y la distancia en que se encuentra hacia la cámara, se tiene en cuenta que mientras más lejos este el rostro de la cámara, la calidad de la imagen disminuye, esto hace que la red neuronal se complique al momento de tratar de detectar un rostro.



Figura 10: Segunda Prueba de detección de rostros.

Para esta segunda prueba se puede observar una mejora con respecto a la imagen anterior, ya que se mejoró con la calidad de la imagen con respecto a la luminosidad, teniendo aun dificultades para detectar un rostro que no este fijo hacia la cámara.

IV. CONCLUSIONES Y RECOMENDACIONES

3.1. Conclusiones

- Se trabajó con el Lenguaje de Programación Python, el cual a través de sus librerías, nos permitió poder acceder a la cámara del dispositivo en donde se desarrolló el software, este fue el primer avance, ya que mediante este procedimiento se tiene acceso a las imágenes que emite la cámara de mismo dispositivo o una cámara externa, de esta manera se pueden realizar las primeras pruebas con respecto a la detección del rostro.
- Inicialmente la red neuronal que se desarrolló permitió detectar de manera correcta los rostros de manera frontal, donde se logró obtener el 100% de aciertos y un falso positivo de 0% sin embargo, se pudo notar que cuando hay una iluminación bastante grande sobre los segmentos de video vigilancia, el rendimiento de la red neuronal baja de manera considerable. Por otro lado, al obtener una rotación en el rostro, la capacidad de la red neuronal para la detección de rostros disminuye, por lo que, si deseamos obtener mejores resultados, se debe implementar una red neuronal más robusta que permita detectar rostros ya sea de manera frontal o con una rotación en el mismo.
- En la siguiente fase de la creación del almacenamiento de los rostros detectados, mostró una cantidad favorable con los rostros detectados en las secuencias de video, ya que se logró obtener un valor diminuta de falsos positivos sobre los rostros que se planteó detectar, esto se debe a que la red neuronal pierde rendimiento por temas de iluminación u otros aspectos, los cuales nos permite generar falsos positivos.
- La compatibilidad que muestra la red neuronal con los distintos lugares de desarrollo y su relativa facilidad para implementarlos al mostrar un lenguaje de programación disponible para estructuras Python, permite el uso de librerías opencv que ayudan al manejo de herramientas de visión artificial y detección facial, de esta manera se pudo obtener resultados altos, en los diferentes escenarios en donde se realizaron las pruebas, obteniendo mejores resultados en una secuencia de video con imágenes frontales sobre una cantidad más baja sobre imágenes con mayor iluminación o diferentes aspectos del rostro.

3.2. Recomendaciones

- Para poder solucionar el problema de la rotación del rostro se recomienda trabajar con las siguientes propuestas:
- Base de datos donde tengan imágenes de un humano en distintas rotaciones del rostro.
- Métodos híbridos, en donde hay distintas imágenes por persona durante el entrenamiento.
- Para el problema de iluminación se recomienda trabajar con los siguientes métodos:
 - a. Métodos Heurísticos.
 - b. Métodos basados en la clase.
 - c. Métodos basados en el modelo 3D.
- Para videos de alto procesamiento computacional se recomienda utilizar una alta memoria GPU, para que el procesamiento de instrucciones y accesos a la memoria se pueda realizar de una manera más veloz y no se tengan limitaciones de almacenamiento de datos y uso de memoria.
- Para la realización de esta aplicación se recomienda trabajar con la última actualización del Opencv, ya que esta cuenta con funcionalidades nuevas y mejoradas para la detección de rostros.

Referencias

- [1] T. S. N. N. Shivalila Hangaragi, «Face Detection and Recognition Using Face Mesh and Deep Neural Network,» *Procedia Computer Science* 218 (2023) 741–749, India, 2023.
- [2] D. K. y. N. A. Ming-Hsuan Yang, «Detecting Faces in Images: A Survey,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- [3] S. D. F. Hernández, «DESARROLLO DE UN ALGORITMO DE DETECCION FACIAL Y SU APLICACIÓN A UN ATRIL AUTOMATIZADO,» Toluca, Mexico, 2022.
- [4] S. M. GONZALEZ, «Implementación y Comparación de dos Algoritmos Supervisados en Redes Neuronales Convolucionales Orientadas a la Detección de Rostros para Ejecutarse en Hardware de bajos recursos,» Universidad Autónoma de Querétaro, Santiago de Querétaro, Querétaro, México, 2021.
- [5] C. V. Alcantara Montiel, J. C. Pedraza Ortega, J. M. Ramos Arreguín y E. Gorrostieta Hurtado, «Detección efectiva de rostros en imágenes utilizando descriptores basados en HOG,» *Research in Computing Science*, Queretaro, Mexico, 2019.
- [6] A. Capcha Huaman y J. L. Huaman Chile, «Sistema de Reconocimiento Facial Basado En Los Algoritmos Haar Cascade, Deep Faces y Luxand FacesSDK,» Lima - Perú, 2021.
- [7] Z. Á. Jeremias Ambrogio, «Reconocimiento de objetos a través de la Metodología Haar Cascade,» *Revista Argentina de Ingenieria*, Córdoba, Argentina, 2020.
- [8] R. Gonzales Lozan, «Sistema de Reconocimiento Facial con Deep Learning», en Leganés, Ecuador nos mencionan que, en los últimos años,» Leganés, España, 2020.
- [9] A. B. C. Menguiano, «Integración y evaluación de sistemas de reidentificación de caras usando MTCNN y FaceNet en C++,» Sevilla,

España, 2019.

- [10] D. M. F. C. P. V. Granja, «Procesamiento de imágenes para la identificación de personas como sistema de seguridad en zonas domiciliarias,» de *VI Congreso Internacional Sectei 2019*, 2019.
- [11] J. F. C. Fonseca, «DISEÑO DE UN SISTEMA DE APOYO EN SEGURIDAD ~,» 2019.
- [12] C. Y. M. Díaz, «Nuevos métodos para el reconocimiento automático de rostros en video,» de *2019 IEEE 39th Central America and Panama Convention (CONCAPAN XXXIX)*, 2019.
- [13] R. J. C. Talavera, «MODELO PARA LA DETECCIÓN DE ROSTROS EN SECUENCIAS DE VIDEO CON BAJA CALIDAD,» Arequipa, Perú, 2018.
- [14] J. E. Salcedo Soto, «Detección y Reconocimiento de Rostros en Imágenes Infrarrojas sobre Hardware Digital Dedicada,» Concepción, Chile, 2016.
- [15] A. Y. Bernal Leyva, «Análisis de Métodos de Reconocimiento Facial Bajo el Sistema Operativo Android,» Pimentel, Perú, 2018.
- [16] Y. Miranda Diaz, E. E. Gallego Martinez y L. Margolles Clausel, «Detección de Rostros Empleando MATLAB,» *REVISTA SARANCE* N° 39, 2017.
- [17] E. R. Caballero Barriga, «Aplicación Práctica de la visión Artificial para el Reconocimiento de Rostros en una Imagen, Utilizando Redes Neuronales y Algoritmos de Reconocimiento de Objetos de la Biblioteca Opencv,» Bogotá, Colombia, 2017.
- [18] Kirti Dang y Shanu Sharma, «Review and Comparison of Face Detection Algorithms,» *International Conference on Cloud Computing, Data Science & Engineering – Confluence*, Noida, India, 2017.
- [19] S. D. F. Hernández, «Desarrollo de un nuevo Algoritmo de Detección Facial y su Aplicación a un Atril Automatizado,» Toluca, 2022.
- [20] E. Á. J. Ambrogio, «Reconocimiento de objetos a través de la

metodología Haar Cascade,» *REVISTA ARGENTINA DE INGENIERÍA*, pp. 65-66, 2020.

- [21] D. C. N. H. G. D. L. C. F. Dr. C. Heydi Méndez Vázquez, «Nuevos métodos para el reconocimiento automático de rostros en video,» Diciembre 2019.
- [22] S. M. C. Leyton, «Deteccion de Deepfakes,» Repositorio Institucional Séneca, Bogotá, Colombia, 2022.
- [23] R. G. Lozano, «Sistema de Reconocimiento Facial con Deep Learning,» Universidad Carlos III de Madrid, Leganés, España, 2020.

ANEXOS



FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO RESOLUCIÓN N°0307-2023/FIAU-USS

Pimentel, 28 de abril de 2023

VISTO:

El Acta de reunión N°02504-2023 del Comité de investigación de la Ingeniería de Sistemas remitida mediante Oficio 0143-2023/FIAU-IS-USS de fecha 26 de abril de 2023, y;

CONSIDERANDO:

Que, de conformidad con la Ley Universitaria N° 30220 en su artículo 48° que a letra dice: "La investigación constituye una función esencial y obligatoria de la universidad, que la fomenta y realiza, respondiendo a través de la producción de conocimiento y desarrollo de tecnologías a las necesidades de la sociedad, con especial énfasis en la realidad nacional. Los docentes, estudiantes y graduados participan en la actividad investigadora en su propia institución o en redes de investigación nacional o internacional, creadas por las instituciones universitarias públicas o privadas.";

Que, de conformidad con el Reglamento de grados y títulos en su artículo 21° señala: "Los temas de trabajo de investigación, trabajo académico y *tesis* son aprobados por el Comité de Investigación y derivados a la facultad o Escuela de Posgrado, según corresponda, para la emisión de la resolución respectiva. El periodo de vigencia de los mismos será de dos años, a partir de su aprobación. En caso un tema perdiera vigencia, el Comité de Investigación evaluará la ampliación de la misma.

Que, de conformidad con el Reglamento de grados y títulos en su artículo 24° señala: La tesis es un estudio que debe denotar rigurosidad metodológica, originalidad, relevancia social, utilidad teórica y/o práctica en el ámbito de la escuela profesional. Para el grado de doctor se requiere una tesis de máxima rigurosidad académica y de carácter original. Es individual para la obtención de un grado; es individual o en pares para obtener un título profesional. Asimismo, en su artículo 25° señala: "El tema debe responder a alguna de las líneas de investigación institucionales de la USS S.A.C."

Que, mediante documentos de vistos, el Comité de investigación de la referida Escuela profesional acordó aprobar la ampliación de la vigencia de las tesis que se detallan en el Acta de reunión N° 02504 - 2023, de la línea de investigación de INFRAESTRUCTURA, TECNOLOGÍA Y AMBIENTE, a cargo de los estudiantes y /o egresados del Programa de estudios INGENIERÍA DE SISTEMAS, hasta la fecha que indica la presente resolución.

Estando a lo expuesto, y en uso de las atribuciones conferidas y de conformidad con las normas y reglamentos vigentes;

SE RESUELVE:

ARTÍCULO 1: AMPLIAR VIGENCIA, de la Tesis a cargo de los estudiantes y /o egresados del Programa de estudios de **INGENIERÍA DE SISTEMAS** que se detallan en el anexo de la presente Resolución.

ARTÍCULO 2: DEJAR SIN EFECTO, toda Resolución emitida por la Facultad que se oponga a la presente Resolución.

Pimentel, 28 de abril de 2023

ANEXO

AMPLIACION DE VIGENCIA DE TEMA DE TESIS

	APELLIDOS	TESIS	OBSERVACION	RESOLUCION APROBACIÓN/A MPLIACIÓN	ESTADO
1	SANTISTEBAN QUISPE MIGUEL ANGEL	Evaluación del rendimiento de Support Vector Machine y Random Forest en ataques denegación de servicio distribuido en la red de área local	Se amplía hasta el 31 de julio del 2023	0643-A-2018/FIAU-USS	APROBADO
2	RIVADENEIRA PURISACA KARINA MERCEDES	Diseño de un modelo de gestión de pruebas de software basado en la norma ISO 29119-2 procesos de prueba aplicado en un caso de estudio	Se amplía hasta el 31 de julio del 2023	0012-A-2022/FIAU-USS	APROBADO
3	MERINO ANCAJIMA JHENSON JHAMPIER	Sistema de reconocimiento de rostros mediante cámaras de seguridad interna	Se amplía hasta el 30 de mayo del 2023	1359A-2015/FIAU-USS	APROBADO
4	RAMIREZ BENDEZU JOSE CARLOS	Evaluación de las técnicas de aseguramiento de la información sensible de una universidad de Perú	Se amplía hasta el 31 de julio del 2023	1865-A-2019/FIAU-USS	APROBADO



USS Dr. Victor Alexei Tuesta Montaña
Decano (E) / Facultad de Ingeniería,
Arquitectura y Urbanismo
UNIVERSIDAD SEÑOR DE SIPÁN SAC.



USS Dr. Halyn Alvarez Vásquez
Secretario Académico Facultad de
Ingeniería, Arquitectura y Urbanismo
UNIVERSIDAD SEÑOR DE SIPÁN SAC.

Cc: Interesado, Archivo

REGÍSTRESE, COMUNÍQUESE Y ARCHÍVESE

Anexo 2:

La implementación del método se realizó en el lenguaje de programación Python, y para el proceso de entrenamiento de imágenes se trabajó con una base de datos de imágenes positivas (imágenes con rostros) y negativas (imágenes sin rostros).

Fase de Entrenamiento

Para la fase de entrenamiento de las imágenes se utilizó la herramienta Cascade Trainer GUI, esta herramienta permite entrenar, testear y mejorar modelos de clasificación de cascada. Básicamente proporciona un mecanismo para obtener un set de datos con imágenes que corresponden al objeto que estamos buscando detectar (imágenes positivas) y otro set con imágenes que no coinciden con dicho objeto (imágenes negativas), a partir de estas fuentes de datos se obtiene como salida (resultado) un modelo de cascada en formato XML.

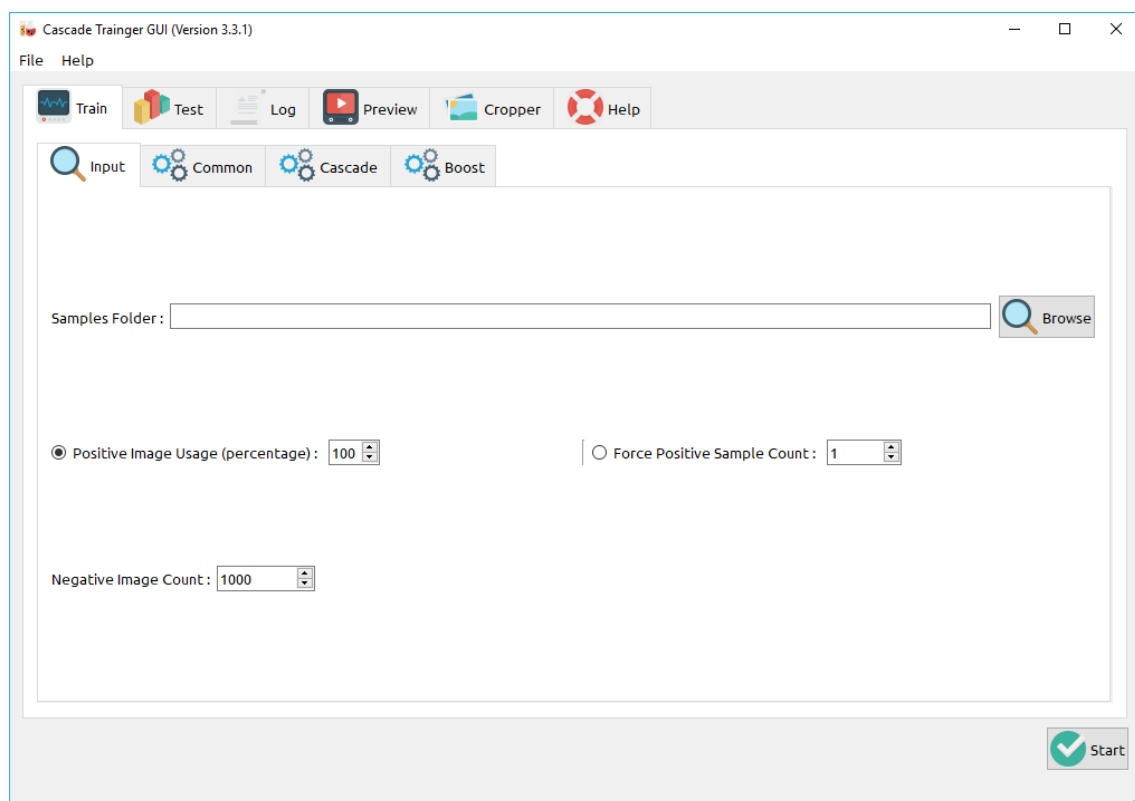


Figura 11: Interfas de entrenamiento.

haarcascade_frontalface_alt 23/03/2023 14:30 Documento XML 899 KB

Figura 12: Archivo XML con la red neuronal entrenada

en esta herramienta es donde se ingresaron las imágenes con un total de 500 imágenes positivas y 500 imágenes negativas, proporcionando el tamaño de estas imágenes en pixeles.



Figura 13: Imágenes Positivas

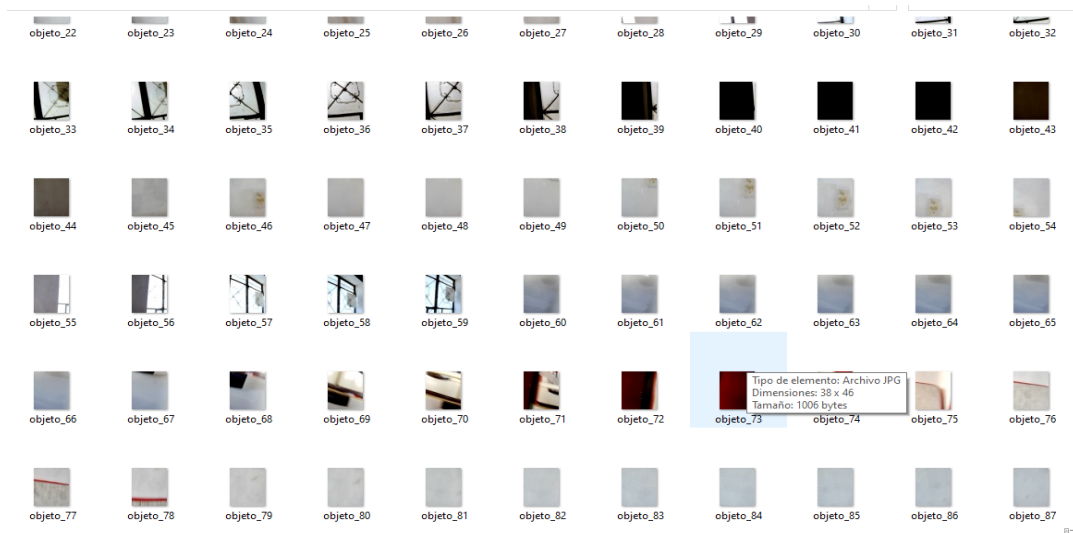


Figura 14: Imágenes Negativas

En el siguiente gráfico se explica la creación de la red neuronal, en donde la imagen de entrada nos muestra una imagen de un rostro en donde cada pixel de la imagen se convierte en una capa que contiene partes de la imagen, cada parte de la imagen pasa a una capa de convolución con el fin de que se pueda extraer las líneas o contornos de la imagen y estas van a la capa de agrupación la cual une todas las

características de la capa de convolución con el fin de que las capas ocultas las puedan clasificar y entregar a la capa de salida con la imagen que se pretende reconocer.

Esquema de la Red Neuronal

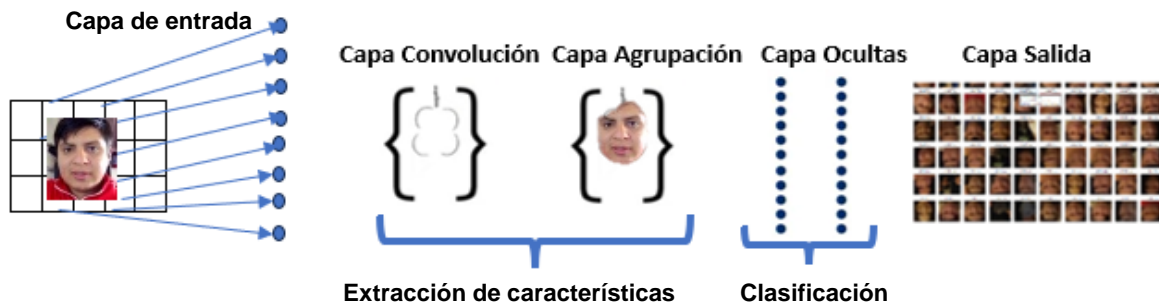


Figura 15: Esquema de la red neuronal.

Zequiél Ángel Jeremías Ambrogio [7], En su investigación “Reconocimiento de objetos a través de la metodología Haar Cascades” propone la utilización de la herramienta Cascade Trainer, en donde es comparado con un proyecto de OpenCv Source, para este modelo de entrenamiento se utilizó 200 imágenes, donde se obtuvieron resultados muy buenos con porcentajes del 94% de precisión.

En comparación con nuestro modelo entrenado los resultados son más elevados, teniendo en cuenta la cantidad de imágenes con las que se ha realizado el entrenamiento, este método es eficaz, pero cuenta con las siguientes limitaciones:

- a. Mientras más imágenes se desee entrenar la red neuronal va a necesitar bastante memoria RAM y memoria de video
- b. La calidad el video tiene que ser buena, ya que al no tener una imagen nítida el entrenamiento no será el adecuado
- c. El tiempo en que demora el entrenamiento depende de la cantidad de imágenes que se desee entrenar.

Instalación de PYTHON

Para la instalación de Python necesariamente se tiene que descargar el archivo de instalación que se encuentra en el siguiente enlace:

<https://www.python.org/downloads/>

Una vez descargado el archivo de Python tenemos que instalar en nuestra Computadora o Laptop como se visualiza en el anexo 1 (Anexo Imagen 1), es de suma importancia instalar Python para poder trabajar con el código de la detección de rostros, para esto vamos a seguir los siguientes pasos:

Paso 1: Ejecutamos el archivo descargado



Figura 16: Paso 1 de la instalación de Python.

Cuando ejecutamos el archivo descargado nos figura esta ventana, en la parte baja seleccionamos Add Python 3.9 to PATH y luego damos click en Install Now.

Paso 2: Esperamos a que termine de cargar el archivo de instalación

Una vez realizado ese proceso nos figura la siguiente ventana, en donde tenemos que esperar que cargue la barra.

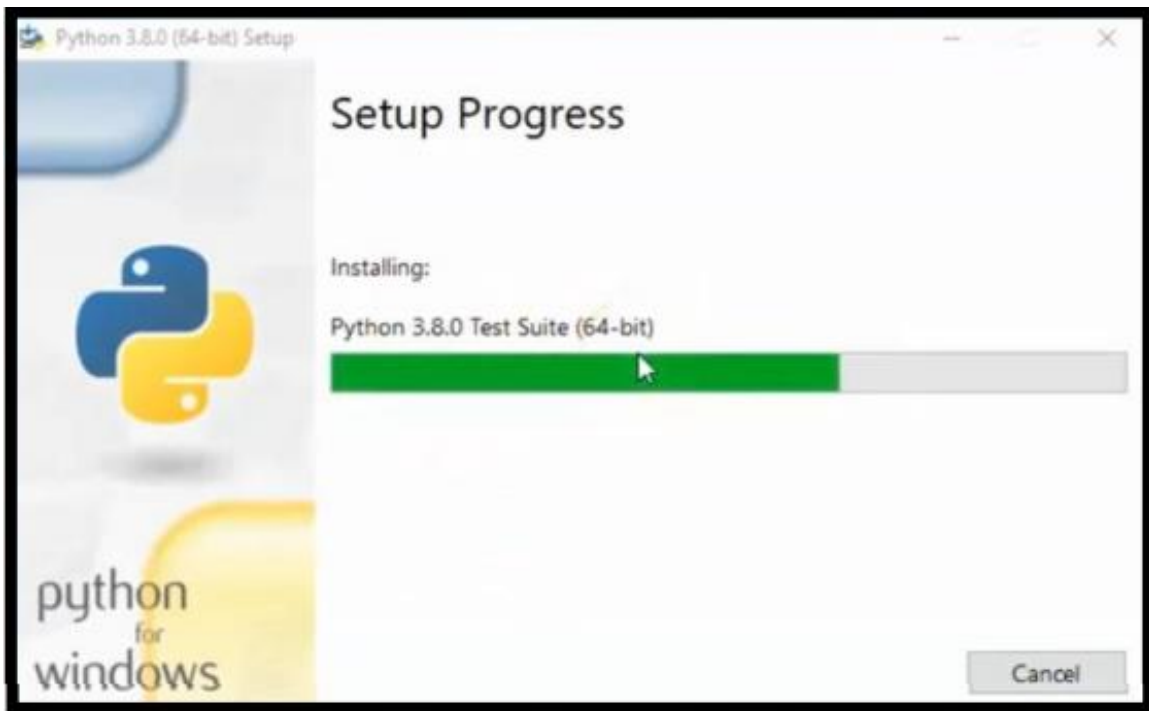


Figura 17: Paso 2 de la instalación de Python.

Paso3: Finalización de instalación

Una vez cargado la barra de instalación, concluimos la fase de instalación de Python.

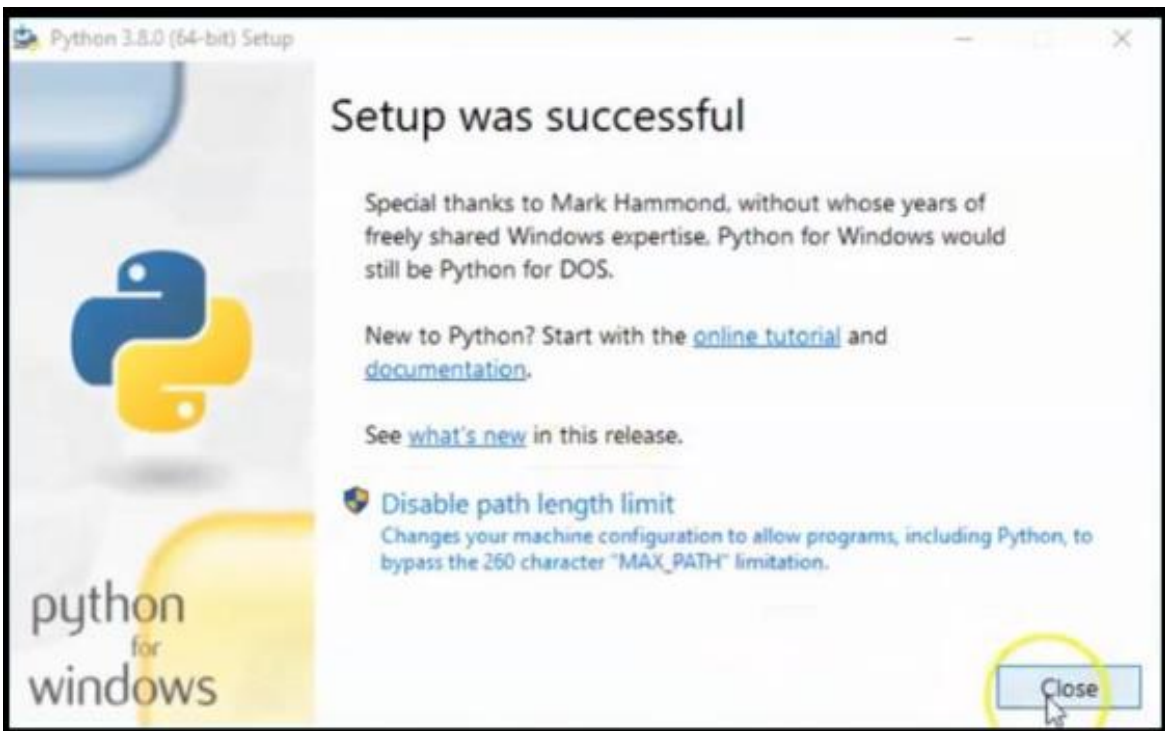


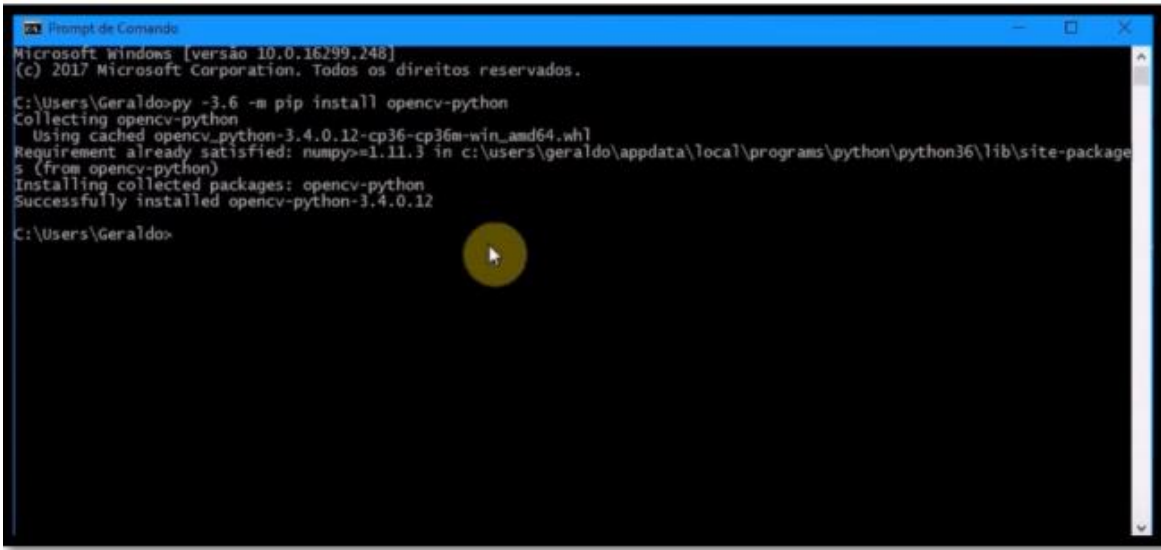
Figura 18: Finalización de instalación de Python.

Anexo 3: Instalación de la librería OpenCv

Una vez instalado Python procedemos a instalar la librería OpenCv para el procesamiento de imágenes, para esto tenemos que instalar las siguientes librerías: Para instalar la librería OpenCv, primero tenemos que abrir el CMD y escribir la siguiente línea de código:

instalar opencv: pip install opencv-python

Tal y como se muestra en la Imagen 14



```
Microsoft Windows [versión 10.0.16299.248]
(c) 2017 Microsoft Corporation. Todos los derechos reservados.

C:\Users\Geraldo>py -3.6 -m pip install opencv-python
Collecting opencv-python
  Using cached opencv_python-3.4.0.12-cp36-cp36m-win_amd64.whl
Requirement already satisfied: numpy>=1.11.3 in c:\users\geraldo\appdata\local\programs\python\python36\lib\site-packages (from opencv-python)
Installing collected packages: opencv-python
Successfully installed opencv-python-3.4.0.12

C:\Users\Geraldo>
```

Figura 19: Instalación de la librería OpenCv.

Una vez realizado la instalación de la librería OpenCv, se procede a instalas las siguientes librerías, realizando el mismo procedimiento:

Instalación de la herramienta imutils:

Para poder instalar la herramienta de imutils, primero tenemos que dirigir al siguiente enlace:

<https://pypi.org/project/imutils/>

una vez dentro, tenemos que copiar el código: pip install imutils, tal y como se observa en la imagen 15:

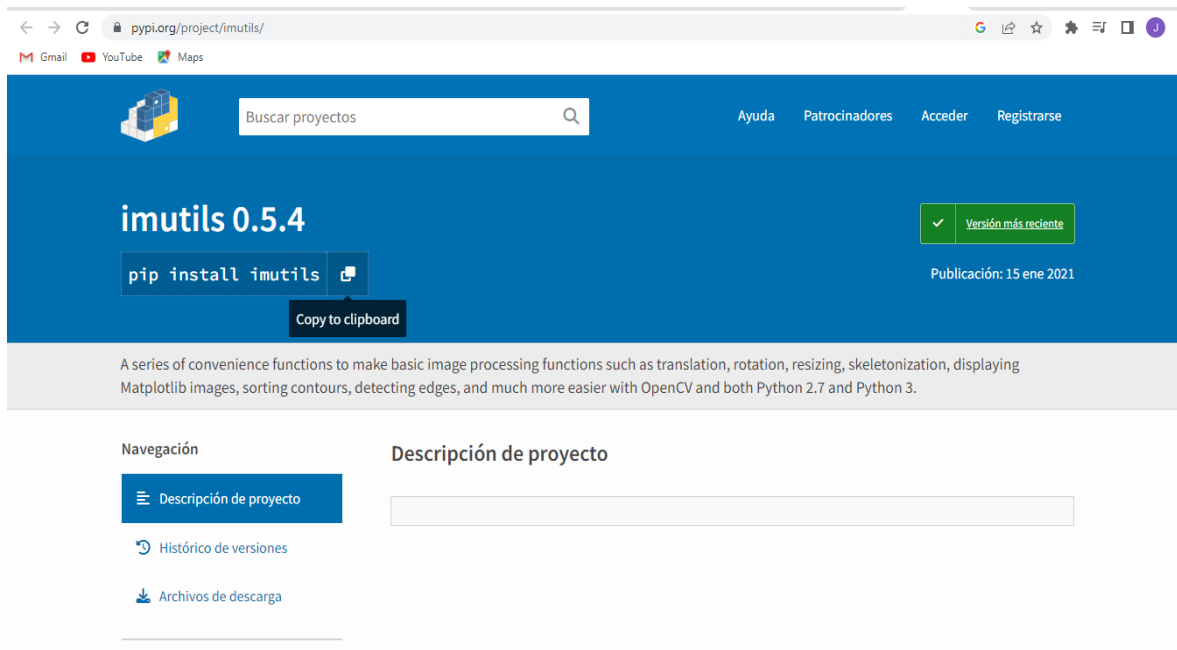


Figura 20: Copiar línea de código imutils.

Esa línea de código la pegamos en el CMD y esperamos que se instale.

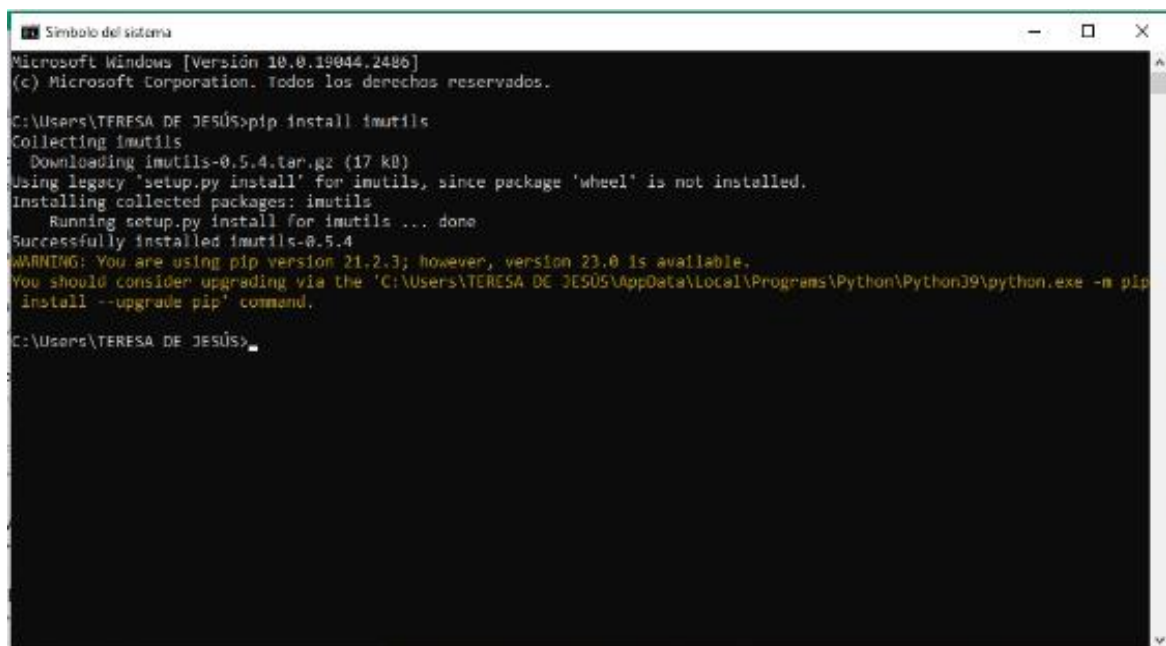


Figura 21: instalación de la librería imutils.

Instalación de la librería MediaPipe:

Para poder instalar la herramienta de mediapipe, primero tenemos que dirigir al siguiente enlace:

<https://pypi.org/project/mediapipe/>

una vez dentro, tenemos que copiar el código: `pip install mediapipe`, tal y como se observa en la imagen 17:

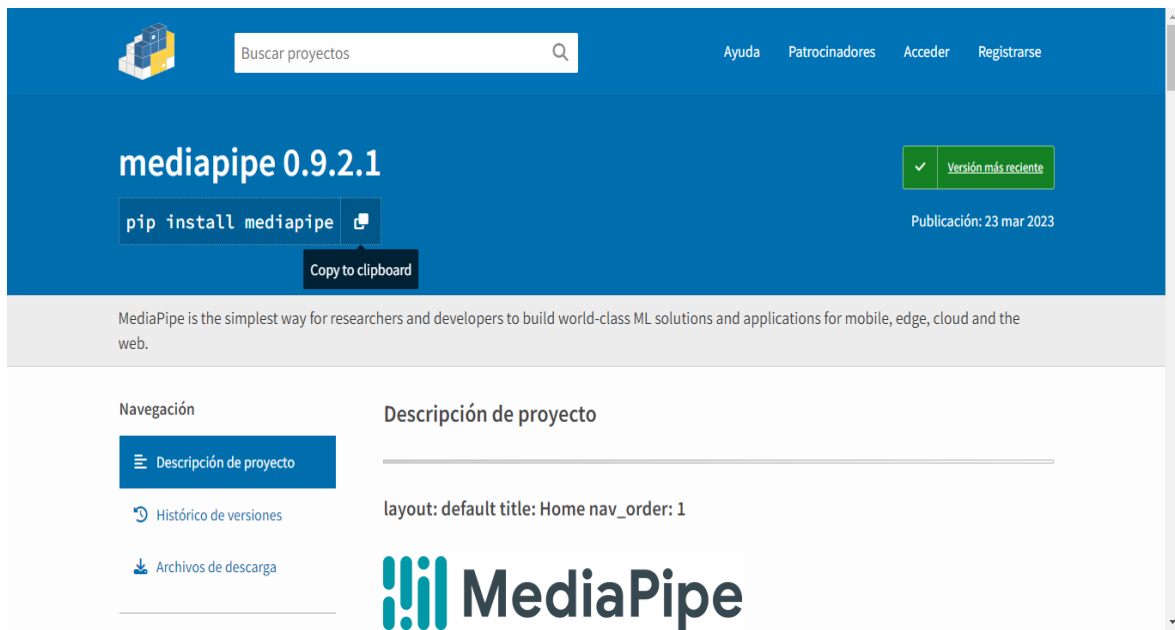


Figura 22: Copiar línea de código de MediaPipe.

Esa línea de código la pegamos en el CMD y esperamos que se instale

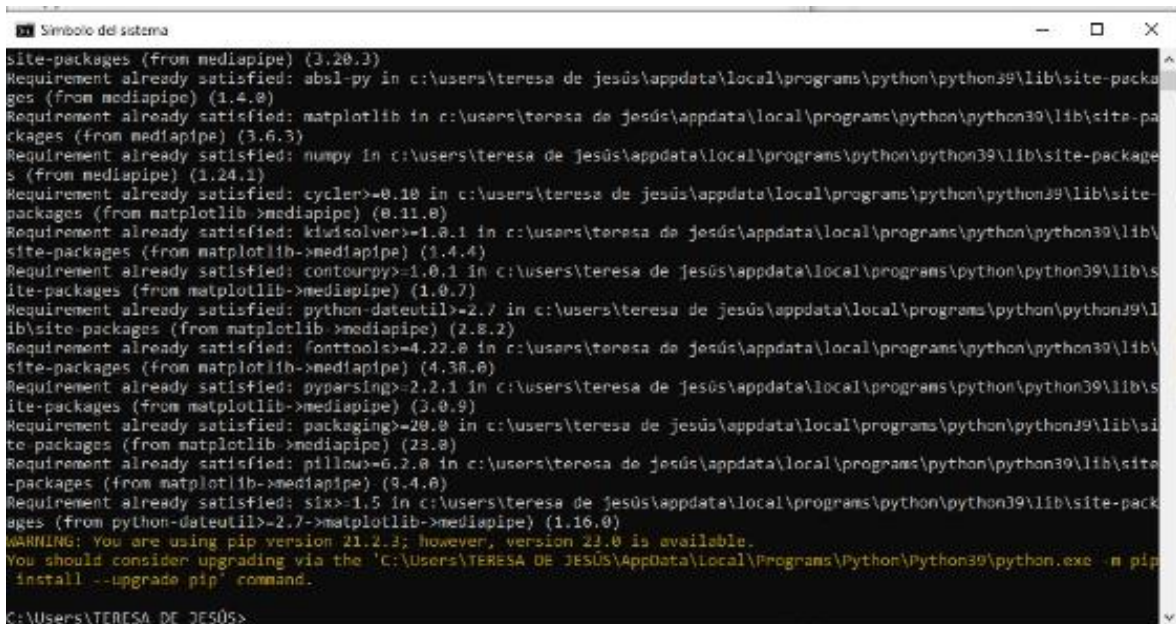


Figura 23: Instalación MediaPipe.

Una vez instalada estas librerías ya se procede a la realización del aplicativo, el cual nos va a permitir detectar rostros en imágenes de videos y fotografías en ambientes semi controlados.

Anexo 4: Interfaz de Detección de rostros

La interfaz se realizó teniendo todas las librerías de OpenCv instaladas en Python, para la realización de este proyecto se utilizó el siguiente código:

```
from tkinter import *
from tkinter import filedialog
import cv2
import imutils
from PIL import Image, ImageTk
import os

video = None
archivo = None
ban = None
image=None

if not os.path.exists('Rostros encontrados'):
    print('CARPETA CREADA')
    os.makedirs('Rostros encontrados')

ventana=Tk()
ventana.title('MULTIMEDIA')
ventana.geometry('900x600')
ventana.configure(bg='beige')

video = None
archivo = None
archivo_mp3 = None
ban = None
image=None
cascada_rostro = cv2.CascadeClassifier('haarcascade_frontalface_alt.xml')
cont=0
auxFrame=None
```

```

coordenadas_rostros=None
def video_stream():
    global video
    global archivo
    global ban
    global cascada_rostro
    if ban==1:
        video=cv2.VideoCapture(archivo)
        iniciar()
    if ban == 0:
        global image
        image = Image.open(archivo)
        image = ImageTk.PhotoImage(image.resize((600,450)))
        iniciar_imagen()

def iniciar_imagen():
    global image
    global cascada_rostro
    lblVideo.place(x=50, y=100)
    lblVideo.configure(image=image)
    lblVideo.image =image
    lblVideo.after(10,iniciar_imagen)
def iniciar():
    global archivo_mp3
    global video
    global cascada_rostro
    global auxFrame
    global coordenadas_rostros
    ret,frame=video.read()
    if ret == True:
        lblVideo.place(x=50, y=100)
        frame = imutils.resize(frame,width=600,height=600)
        auxFrame = frame.copy()

```



```

img_gris = cv2.cvtColor(frame,cv2.COLOR_BGR2GRAY)
coordenadas_rostros = cascada_rostro.detectMultiScale(img_gris, 1.1,1)

for (x,y,ancho,alto) in coordenadas_rostros:
    cv2.rectangle(frame,(x,y),(x+ancho,y+alto),(0,0,255),3)
frame = cv2.cvtColor(frame,cv2.COLOR_BGR2RGB)
img = Image.fromarray(frame)
image = ImageTk.PhotoImage(image=img)
lblVideo.configure(image=image)
lblVideo.image =image
lblVideo.after(10,iniciar)
else:
    video.release()

def finalizar():
    global video
    global ban
    global image
    if ban==1:
        lblVideo.place_forget()
        video.release()
    if ban==0:
        lblVideo.place_forget()
        image=None

def abrirArchivo():
    global archivo
    global archivo_mp3
    global ban
    archivo = filedialog.askopenfilename(title='abrir',initialdir='C:/Users/TERESA/Documents/tesis')
    ban=1

```

```

def abrirlImagen():
    global archivo
    global archivo_mp3
    global ban
    ban=0
    archivo = filedialog.askopenfilename(title='abrir',initialdir='C:/Users/TERESA/Documents/tesis')

def guardarImagen():
    global cascada_rostro
    global video
    global cont
    global auxFrame
    global coordenadas_rostros
    for (x,y,ancho,alto) in coordenadas_rostros:
        rostro = auxFrame[y:y+alto,x:x+ancho]
        rostro = cv2.resize(rostro,(150,150),interpolation=cv2.INTER_CUBIC)
        cv2.imwrite('Rostros encontrados/rostro_{}.jpg'.format(cont),rostro)
        cv2.imshow('rostro',rostro)
        cont=cont+1
    lbl=Label(ventana,text='BIENVENIDO',bg='beige',font=('Arial',13),fg='blue')
    lbl.place(x=420,y=20)

    lbl=Label(ventana,text='VIDEO',bg='gray',font=('Arial',13),fg='black',width=66,height=1)
    lbl.place(x=50,y=80)

    btn = Button(ventana, text='Abrir archivo',bg='gray',command=abrirArchivo,width=15)
    btn.place(x=700,y=140)

    btn_imagen = Button(ventana, text='Abrir imagen',bg='gray',command=abrirlImagen,width=15)
    btn_imagen.place(x=700,y=210)

    lblFondo=Label(ventana,bg='black')
    lblFondo.place(x=50,y=100,width=600,height=450)

```

```
lblVideo=Label(ventana,bg='black')
```

```
lblVideo.place(x=50,y=100)
```

```
inicio=Button(ventana,text='Iniciar',bg='gray',fg='black',command=video_stream,width=15)
```

```
inicio.place(x=700,y=350)
```

```
fin=Button(ventana,text='Finalizar',bg='gray',fg='black',command=finalizar,width=15)
```

```
fin.place(x=700,y=420)
```

```
btn_guardar_imagenes = Button(ventana,text='Guardar  
rostros',bg='gray',fg='black',command=guardarImagen,width=15)
```

```
btn_guardar_imagenes.place(x=700,y=280)
```

```
ventana.mainloop()
```

Una vez terminado el código, la interfaz y el funcionamiento de nuestro proyecto queda de la siguiente manera:



Figura 24: Interfaz gráfica para la detección de rostros.

En esta interfaz podemos detectar los rostros tanto en imágenes fotográficas como en secuencia de videos, como se aprecia en las siguientes imágenes:



Figura 25: Detección de rostros en imágenes fotográficas.



Figura 26: Detección de rostros en imágenes de video.

En donde todos los rostros detectados quedan almacenados en una nueva base de datos así como se muestra en la figura 22:

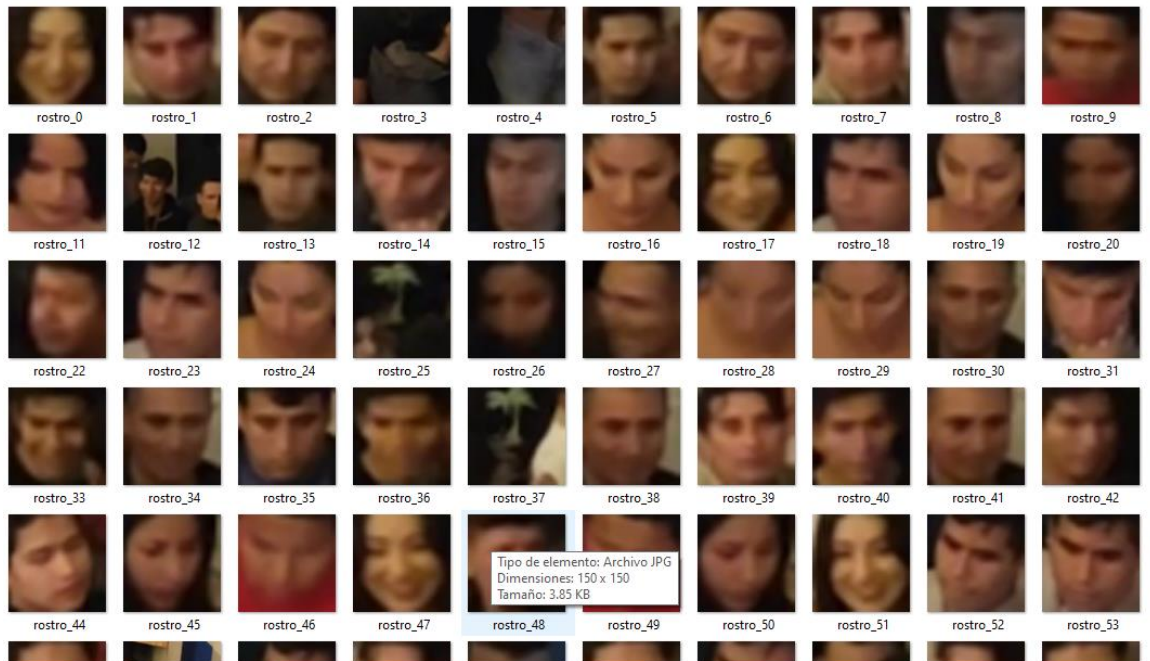


Figura 27: Data base de rostros detectados.

MODELO DE CARTA DE AUTORIZACIÓN PARA EL RECOLECCIÓN DE LA INFORMACIÓN

San Ignacio, 31 de Marzo de 2023

Quien suscribe:

Sr. Careca Santa Cruz Carrasco

Representante Legal – Empresa Gerente MicroRed San Ignacio.

AUTORIZA: Permiso para recojo de información pertinente en función del proyecto de investigación, denominado **DETECCIÓN DE ROSTROS MEDIANTE CAMARAS DE SEGURIDAD INTERNA.**

Por el presente, el que suscribe, **M.C. Careca Santa Cruz Carrasco**, representante legal de la empresa **MicroRed San Ignacio**, AUTORIZO al estudiante(s) **Jhenson Jhampier Merino Ancajima** identificado con DNI N° **70095518**, estudiante del Programa de Estudios de **Ingeniería de Sistemas**, y autor del trabajo de investigación denominado **DETECCIÓN DE ROSTROS MEDIANTE CAMARAS DE SEGURIDAD INTERNA**, al uso de dicha información que conforma el expediente técnico así como hojas de memorias, cálculos entre otros como planos para efectos exclusivamente académicos de la elaboración de tesis, enunciada líneas arriba de quien solicita se garantice la absoluta confidencialidad de la información solicitada.

Atentamente.

RIS GOBIERNO REGIONAL DE CAMBAYA
DIRECCIÓN REGIONAL DE SALUD
RED INTEGRADA DE SALUD SAN IGNACIO
ZONA SANITARIA N°2 SAN IGNACIO
CENTRO DE SALUD SAN IGNACIO
Red Integrada de
Salud San Ignacio

M.C. CARECA SANTA CRUZ CARRASCO CNP N° 81674
GERENTE MICRO RED SAN IGNACIO
Nombre y Apellidos: **Careca Santa Cruz Carrasco**

DNI N°: **45546980**

Cargo de la empresa: **Gerente MicroRed San Ignacio**