



Aplicación de Visión Artificial en Sistemas de Video Vigilancia con Reconocimiento Facial para el Control de Acceso

Daniel Fernando Quinde Ordoñez ^a, Elisa Amelia Cisneros Prieto ^b e Ingrid Angelina Soto Galarza ^c

a. Facultad de Innovación y Desarrollo Tecnológico, Universidad Del Pacífico, daniel.quinde@upacifico.edu.ec

b. Facultad de Innovación y Desarrollo Tecnológico, Universidad Del Pacífico, elisa.cisneros@upacifico.edu.ec

c. Facultad de Negocios, Universidad Del Pacífico, ingrid.soto@upacifico.edu.ec

Palabras Clave

Inteligencia Artificial
Aprendizaje
Algoritmos
Visión Artificial
Estimación

Historia del Artículo

Recibido 04-11-2021

Revisado 05-11-2021

Aceptado 08-11-2021

Publicado 28-12-2021

Resumen. La inmersión de la tecnología en las áreas de la IA, Visión Artificial, Visualización de Datos es un gran reto en la que se deben enfrentar varios aspectos del ambiente real actual, parte de ellos se encuentra la iluminación variable, los entornos dinámicos, la obstrucción parcial de imagen entre varios aspectos.

En la actualidad existen nuevas modalidades de procesamiento y esquemas de clasificación de imágenes para la creación de nuevos modelos de aprendizaje que permitan la identificación y detección en imágenes tanto de acciones como de objetos que formen parte de la captura. Estos modelos operan en base a características cuyo objetivo es culminar con resultados aplicables y útiles.

Dentro de los actuales modelos de técnicas de Aprendizaje de IA, podemos encontrar la red que implica el uso variado con diversos factores. Este modelo específico de redes artificiales promete y consigue muy buenos resultados y son utilizados en investigaciones, uno de los motivos importantes es el hardware que necesita esta tecnología, el cual es muy elevado pero debido a los avances en el campo tecnológico actual permitieron el adelanto progresivo de técnicas de Aprendizaje profundo.

Dichos algoritmos se desarrollan sistemas de altos niveles de confiabilidad, minimizando las probabilidades de obtener caídas del sistema, de manera especial en sistemas de video vigilancia en el cual las máquinas y dispositivos de captura son utilizados, y dentro de los cuales las plataformas de software de identificación y captación de imágenes son lamentablemente de baja calidad cuando no tienen el uso de esta tecnología.

Keywords

Artificial intelligence
Learning
Algorithms
Artificial vision
Estimation

Article History

Recibido 04-11-2021
Revisado 05-11-2021
Aceptado 08-11-2021
Publicado 28-12-2021

Abstract. The immersion of technology in the areas of AI, Artificial Vision, Data Visualization is a great challenge in which several aspects of the current real environment must be faced, part of them is variable lighting, dynamic environments, obstruction partial image between various aspects.

Currently there are new processing modalities and image classification schemes for the creation of new learning models that allow the identification and detection in images of both actions and objects that are part of the capture. These models operate on the basis of characteristics whose objective is to culminate with applicable and useful results.

Within the current models of AI Learning techniques, we can find the network that implies the varied use with various factors. This specific model of artificial networks promises and achieves very good results and they are used in research, one of the important reasons is the hardware that this technology needs, which is very high but due to advances in the current technological field they allowed progressive advancement of Deep Learning techniques.

These algorithms develop systems with high levels of reliability, minimizing the probabilities of obtaining system drops, especially in video surveillance systems in which the capture machines and devices are used, and within which the software platforms of Identification and image capture are unfortunately of low quality when they do not have the use of this technology.

1. Introducción

Como indica el artículo, La Inteligencia Artificial es la virtud de las máquinas para utilizar algoritmos, estudiar los datos y usar lo que se ha aprendido en la toma de decisiones tal y como lo pudiera hacer una persona. Pero, los dispositivos basados en Inteligencia Artificial no necesitan parar a descansar y además analizan gran cantidad de información al mismo tiempo. De esta forma, si se realiza una evaluación de los errores, es menor en las máquinas que ejecutan las mismas tareas que las personas, (Rouhiainen, 2018).

Como se indica en la publicación de la revista científica Quartz: El Deep Learning o también conocido como, el aprendizaje profundo, se logra gracias a la aplicación de redes neuronales, las cuales se distribuyen en capas con la finalidad de reconocer relaciones y patrones complejos que se encuentran inmersos en los datos. Para lograr esto es necesario contar con un gran volumen de información útil para el sistema y una gran capacidad de procesamiento computacional. En la actualidad se maneja sistemas de reconocimiento de voz, así como el procesamiento del lenguaje natural, la aplicación de visión artificial y métodos de identificación de vehículos en los sistemas de asistencia al conductor, (Gershgorin, 10 de septiembre de 2017).

Como se indica en la Revista Innovación, se ha logrado mejorar el desarrollo de soluciones que se basan en el aprendizaje que logran las máquinas gracias a los algoritmos avanzados que le permiten emular los niveles de análisis similares al funcionamiento del cerebro humano. (Innovación, 2019).

La video vigilancia, son un grupo de análisis de formas de comportarse de alta especialización, que sirven para precisar acciones propias que se pueden alcanzar en el momento de visualizar las cámaras. (Contactcenteronline, 2019).

Los sistemas de video vigilancia se encuentran en una mejora continua por lo tanto el reconocimiento facial y la inteligencia artificial en general son primordiales para generar nuevas tecnologías que ayudan a controlar los niveles de seguridad de las personas. (Contactcenteronline, 2019).

Los dispositivos que utilizan el deep learning, se aplican desde temas básicos como eliminar prescindir de falsos positivos en los sistemas de video grabación y vigilancia, hasta algo más complejo como el reconocimiento del objeto en la escena que se analiza. (Contactcenteronline, 2019).

La tecnología de aprendizaje en profundidad ha sido aplicada para el reconocimiento de personas basándose en aspectos tanto físicos como comportamiento humano con la finalidad de ubicar a fugitivos de la justicia, así como también ayudar a encontrar a personas desaparecidas y aplicar esta tecnología para prevenir posibles delitos o detectar infracciones en estacionamientos. (Contactcenteronline, 2019).

Según el autor Culjak, del artículo científico relacionado destaca que en concepto: La visión artificial permite programar un equipo con el algoritmo idóneo para procesar tanto imágenes como videos y llegar al punto de lograr interpretarlos. (Culjak & Abram, D, 2013). Dicho proceso entonces se basa en la transformación de datos cuyo origen puede ser un video o fotograma de cámara representado en una posible decisión o alguna nueva característica.

2. Marco Teórico

Los sistemas de reconocimiento facial han venido evolucionando y continuamente mejorando en los veinte últimos años, su principal razón, es que son ampliamente utilizados en algunos objetos de control. Como ejemplo tenemos los sistemas biométricos, cuyo objetivo principal es la identificación del individuo a través de una imagen digital de forma automática.

Estos sistemas tuvieron sus inicios en los años 60. Uno de los primeros sistemas era semiautomático y se basaba en la localización de rasgos como ojos, cejas, nariz, entre otras características, para calcular puntos de referencia que se utilizaban para ser analizados con datos almacenados que se encontraban registrados en el sistema.

Las nuevas versiones de reconocimiento facial usaron otros marcadores y mejores técnicas en su proceso de detección. Actualmente existen herramientas de este campo de aplicación que son compatibles en áreas heterogéneas modernas, como son tarjetas inteligentes, controles de acceso, licencias de conducción, identificaciones de pasaportes, acceso a centros de entretenimiento, seguridad y vigilancia, entre otros.

El reconocimiento facial dio un salto para estar dentro de nuestras tareas cotidianas, que actualmente está inmersa en muchas funcionalidades, la implementación de un software implica tareas que son un reto y que requieren de muchas investigaciones que ayuden a mejorar los procesos. Tal como indica la autora María García en su publicación, el desarrollo de los programas y algoritmos necesarios para las plataformas de software de reconocimiento y detección facial usando inteligencia artificial se ha dividido en los siguientes campos: detección e identificación de rostros de personas, seguimiento de los objetos en movimiento, alineamiento del objeto en la imagen, extracción de rasgos faciales e identificaciones, entrenamiento del algoritmo, identificación de objetos, categorización en base a características, investigación de expresión facial, examinación en dos y tres dimensiones (García Chang, 2009)

Como resalta el Autor Andrew Gebhart, en su publicación de la revista científica CNT: El funcionamiento básico de los sistemas de reconocimiento facial consiste en capturar una imagen en dos y tres dimensiones de la cara que se desea evaluar y luego comparar los puntos clave de la imagen con una base de datos de similares características. Aplicando esto a nivel policial, se puede utilizar la fotografía de los delincuentes como la base de datos y en cuanto a cámaras de hogar las imágenes pertenecerían a familiares o amigos que han sido registrados por los sistemas (Gebhart, 2019).

El Autor Gian Luca Foresti, indica que un sistema de video vigilancia se conforma por sensores que se ubican de forma estratégica en el sector que se desea vigilar y están conectados por medio de sistemas de transmisión adecuados en relación con el receptor. (Foresti, 2000).

Este tipo de plataforma de software se basa en dispositivos sensoriales activos, es decir que actúan con el entorno. Y sensores pasivos que no interfieren en el entorno para adquirir los datos. Para el uso de estos sensores se establecen algoritmos que tengan ciertas características para poder cumplir con los criterios respectivos del análisis de información.

Para la creación de estos algoritmos se abarcan conceptos relacionados a la detección e identificación de movimiento registrado, el cual es una de las características principales

en todos los sistemas de video vigilancia. También se tienen métodos para la estimación de las distancias en la que intervienen características como sombras, contornos, texturas, enfoque entre otros factores.

El bloque de mayor importancia en un sistema de procesamiento de imágenes en general, y en especial de video vigilancia, como lo indica el autor Xavier Muñoz Pujol: “El objetivo de la segmentación de imágenes es detectar y extraer las regiones que componen una imagen. Nótese que, contrariamente al problema de clasificación, el reconocimiento de estas regiones no es obligatorio. De hecho, aunque es difícil de concebir, podemos pensar en la segmentación de imágenes como la primera mirada que hicimos al mundo cuando estábamos recién nacido. En otras palabras, mirar sin un conocimiento superior sobre los objetos que podemos ver en la escena. Por tanto, no es posible identificar los diferentes objetos, simplemente porque es la primera vez que los vemos. Entonces, la respuesta de la imagen. El proceso de segmentación será algo como: "hay cuatro regiones en la imagen" y una matriz del tamaño de la imagen donde cada píxel está etiquetado con el correspondiente número de región” (Muñoz, 2003).

En la **Figura 1**, como indica el autor Pedro Gil Jiménez en su artículo científico, destaca que se da a conocer el flujo que se realiza habitualmente en el área de la visión de inteligencia artificial, donde se indica el orden de las fases el cual involucra desde la recolección de la imagen desde el proceso de la misma hasta la realización de un evento en base a los parámetros recolectados y datos obtenidos.

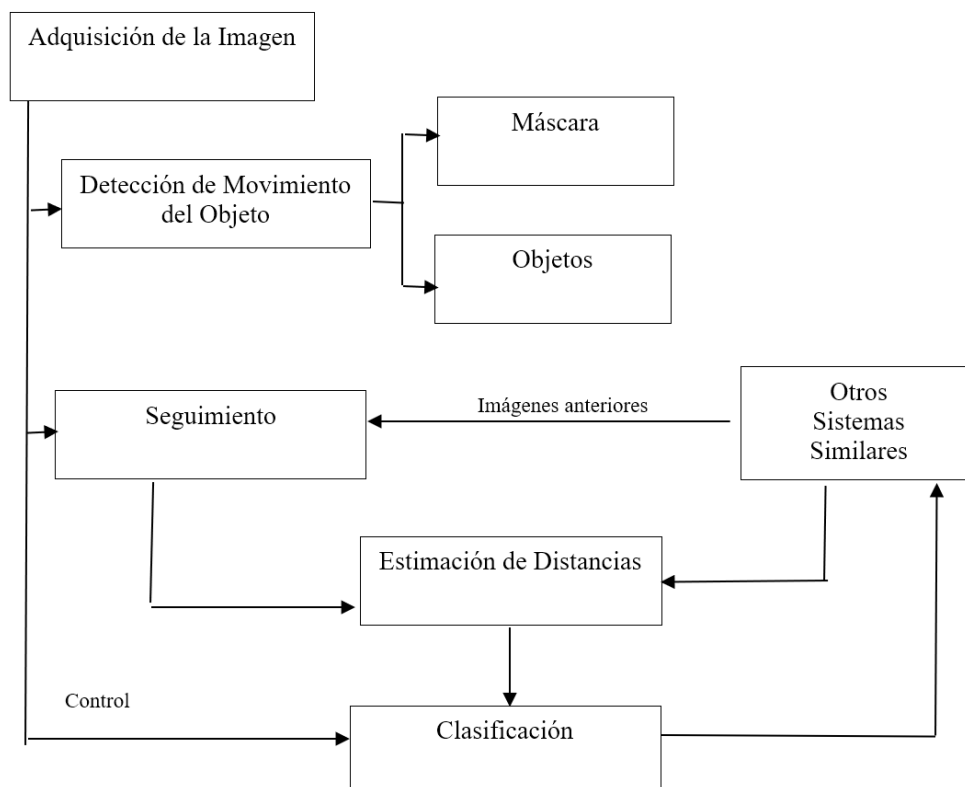


Figura 1. – Gráfica general de procesos de un modelo sistemático de video vigilancia (Jiménez, 2009)

2.1. Detección activa de movimiento mediante Inteligencia Artificial

En función que el procesamiento de vídeo logre realizar los objetivos de identificación y detección de eventos inusuales y extraños se debe incluir código de algoritmos de procesamiento de visión y video artificial que son elaborados específicamente para este propósito. Aunque actualmente se pueden encontrar numerosas aplicaciones de video vigilancia, la gran mayoría de las aplicaciones generalmente incluyen las siguientes características en funciones principales:

Detección de movimiento: Es la función básica del sistema de video vigilancia, dado que la aplicación se enfoca en vigilar las actividades que se generen en un espacio específico, dentro de la implementación de esta solución los sensores de captación se encuentran de forma estática en estructuras inmóviles como un poste desde donde se captan las imágenes. (Gil Jiménez, 2009).

Estimación de distancia: Adicional a determinar que existe movimiento en la escena, también es importante establecer la distancia en la que se encuentran los objetos con respecto al sensor de movimiento. Tal como lo realiza la visión humana y de ciertos animales, es factible estimar la distancia con respecto al sensor gracias al procesamiento de una imagen estereoscópica la cual es captada gracias a varios sensores que se encuentren instalados en la misma escena, pero con cierta separación establecida entre ellas. (Gil Jiménez, 2009).

Cálculo de desplazamiento de cámara: En aplicaciones de video vigilancia, aunque las cámaras son fijas, ellas tienen la capacidad de girar en función de su propio eje gracias a un sistema electromecánico, gracias a esto puede analizar un campo de visión de mayor amplitud que el que se obtendría si el sensor estuviera enfocando un ángulo fijo. El sistema podría conocer durante todo el proceso la posición exacta de la cámara o por lo menos el desplazamiento que existe entre una imagen y la próxima toma. (Gil Jiménez, 2009).

Composición de mosaicos: Cuando es necesario ampliar el área de visión enfocado por una sola cámara, una solución es implementar un sistema con varias cámaras de video con la finalidad que cada una de ellas capturen zonas distintas del sector que se encuentra en vigilancia. Al combinar todas las imágenes es factible formar una imagen única la cual equivale al enfoque obtenido por una cámara con un ángulo de visión mucho mayor. (Gil Jiménez, 2009).

Seguimiento de objetos: Las funciones anteriores se enfocan en determinar la presencia de situaciones sospechosas en la escena, cabe recalcar que en los sistemas de vigilancia es importante la evaluación en tiempo real de las novedades ejecutadas en el sector evaluado, por lo tanto, se enfoca en los objetos extraños que detecta en cada momento. Para lo cual es necesario incorporar funciones que permitan relacionar los objetos en secuencias de imágenes lo cual se denomina seguimiento de objetos. (Gil Jiménez, 2009).

Clasificación de eventos: El objetivo del operador es analizar el evento ocurrido y clasificarlo en base a las experiencias vividas o conocimientos que han sido adquiridos de diferentes formas por la persona que administra el sistema. Los sistemas de segunda generación deben tener incorporada estas funciones y se debe basar en la información generada por las funciones anteriores. Esta funcionalidad maneja diversos objetivos, como determinar si el evento es o no de interés, basándose en ciertos parámetros del objeto detectado, los cuales pueden estar en fusión del comportamiento, ubicación, tamaño o hasta el tipo de objeto que se está evaluando. (Gil Jiménez, 2009).



Figura 2. - Imágenes de predicción izquierda y derecha del fondo estimado (Gil Jiménez, 2009)



Figura 3. – Comparativa de Imágenes con objetos en primer plano (Gil Jiménez, 2009)

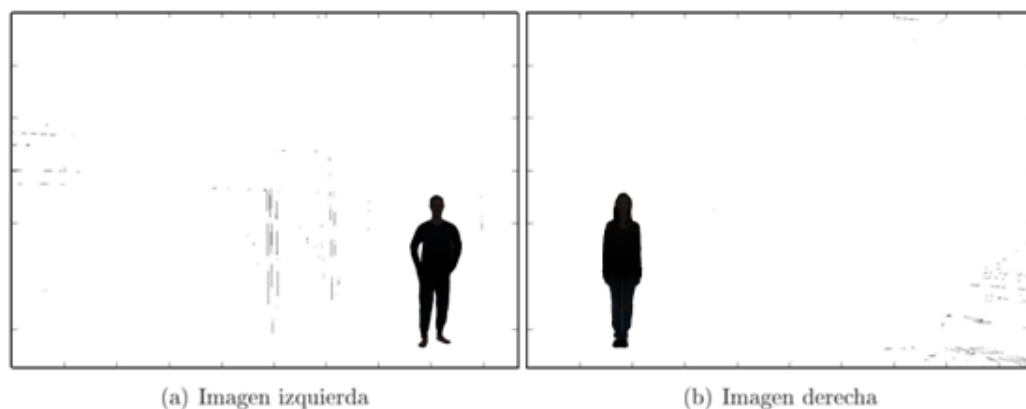


Figura 4. - Máscaras de movimiento para los objetos de la Figura 3 (Gil Jiménez, 2009)

Por otro lado, antes de tomar alguna información que provenga de una imagen, hay que tomar en cuenta las deficiencias que provienen de diferentes fuentes, como indica el investigador y autor John Jairo Sanabria Sarmiento, es posible encontrar (Sanabria, 2011):

- Fallas en el modelo de sistema, lo que seguramente originará defectos en las imágenes.
- Pérdida de datos en la captura.
- El daño de información en los resultados de contingente.
- Las diferentes distorsiones en el escalado de imágenes, las posibles variaciones o quizás fallas obtenidas por la luminosidad disponible.

Por eso, la investigación del autor indica que es necesario el acondicionamiento de las imágenes. Luego se precisan los criterios a destacar involucrando técnicas para optimizarlos como lo es (Sanabria, 2011):

- Redistribuir la información de datos estadísticos de la fuente de imagen.
- Manipular las particularidades del área de coloración (Pureza, Saturación e Intensidad).
- Filtrar la imagen.

Como una fase previa a la clasificación de la información recopilada, es recomendado que se realice el proceso de transformar los datos de la fuente de la imagen a un campo manejable óptimo con el correspondiente compromiso que posteriormente se va a desarrollar.

Es sustancial que se realice la depuración y eliminación que contenga información redundante para un mejor desempeño, conociendo que esto no implique pérdidas relacionadas a información que pueda ser de provecho, por eso es recomendada el uso del modelo conocido de Análisis basado en Componentes Principales (PCA).

El autor de su artículo científico acerca del concepto PCA, Manuel Terrádez Gurrea, indica que el Análisis de Componentes Principales (PCA), es una sólida técnica basada en estadística de sumario de la información, o también conocida por utilizar el proceso de reducción de la dimensión (definida como el número de variables que aportan a la solución). Queriendo indicar que, ante un volumen de datos que tenga una gran cantidad de variables, el propósito será reducir las variables a un número menor considerando que se pierda la menor cuantía de información viable. Los más recientes componentes principales o conocidos como elementos serán parte de una combinación lineal de dichas variables iniciales, y de igual manera serán consideradas independientes entre ellas” (Terrádez Gurrea, 2000).

Tomando apoyo en la lógica del proceso, resaltado por el autor John Sanabria, en el registro y captura de de n datos de muestras, incorporadas mediante variables, se cumple las siguientes condiciones:

$$m \leq n; \quad t \leq \min \{n, m\}$$

Ahora bien, con el propósito de disminuir la cantidad de información que se va a procesar, esperando la reducción de resultados se consiguen los t principales componentes.

Una vez alcanzados los t principales componentes, se ejecuta una combinación lineal aplicada a las variables iniciales que es incorporada por la ecuación siguiente, consiguiendo que las mismas sean variables independientes entre sí (Sanabria, 2011).

$$x = \sum_{a=1}^1 taPa + E$$

Como sostiene el autor R. González, encontrando que ta proviene de la definición de los pesos con datos siendo tratada la relación existente de la información de las muestras tomadas. La variable Pa es considerada que informa de la misma manera la relación existente con las variables declaradas y el conjunto de vectores matriz E procedente al dato de error ocasionado (Gonzalez, 2004).

El proceso de descomposición se ejecuta a través de los vectores propios definidos que se encuentran encontrados en el conjunto de vectores matriz declarada de covarianza que son calculados COV (Sanabria, 2011).

$$COV(X) = \frac{X^t X}{n - 1}$$

$$COV(X)Pa = YaPa \sum_{a=1}^m Ya = 1$$

Siendo la variable Ya , la definición del valor que está asociado al valor propio declarado como Pa . De esta forma puede obtenerse la representación de los datos, que es definida mediante la siguiente fórmula (Sanabria, 2011):

$$t_a = XP_a$$

Basados en la información resultante, con los nuevos datos, se puede identificar que cerca del noventa por ciento (90%) de los datos que se encuentran presentes en los datos de imagen que está asociada en los (3) tres resultados iniciales de componentes principales (Sanabria, 2011).

De forma complementaria al presente análisis realizado con el modelo PCA, se adicionan igualmente modelos de sistemas, basados en conocimiento conocidos como KBS para ejecutar el proceso de clasificación, segmentación y análisis del registro de datos contenido en la fuente de imagen a partir de características.

Como indica el autor Akerkar, acerca del concepto de KBS, resalta que, un modelo de sistema que está basado en el conocimiento (KBS) es un sistema que utiliza técnicas de modelos de inteligencia artificial para resolver los procesos de solución de problemas para

apoyar el sumario de aprendizaje, la toma y selección de decisiones, y la acción humana (Akerkar, R. & Sajja, P., 2009).

3. Metodología

Siendo el propósito final analizar la necesidad evaluándola y analizando la problemática presente, relacionado al procesado y filtrado de imágenes utilizando modelos de visión artificial en las plataformas de software de video vigilancia, se realizó la comparativa partiendo desde los índices de precisión, teniendo en cuenta que la información fue obtenida desde dispositivos de cámaras de seguridad. La precisión obtenida alcanzo un 90% y el tiempo de ejecución alcanza 8 ms por fotograma.

El modelo se base en tres fases, las cuales fueron:

Fase 1: Detección del objeto

Fase 2: Seguimiento del objeto para guardar su posición, se realizó con intervalos de 30 segundos, produciendo una existencia de rastreo por cada objeto.

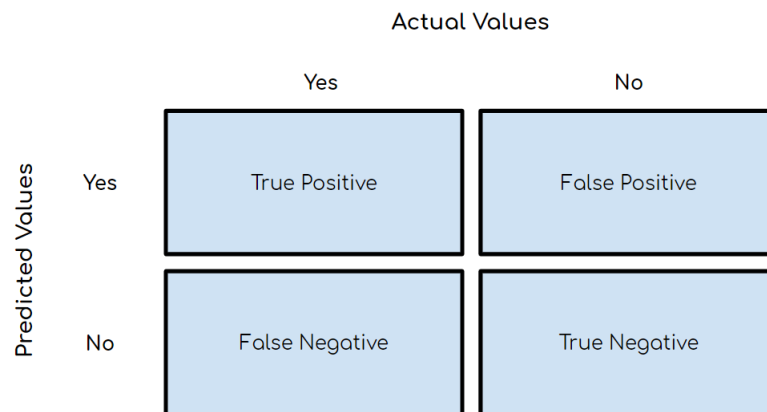
La técnica metodológica de identificación y reconocimiento de vehículos que es utilizado en el tipo propuesto de su tesis está compuesto de filtros de análisis correlacionado de relevancia, que es parte de la propuesta del autor Martin Danelljan, el modelo de código de algoritmo está basado en la evaluación del deslizamiento y escalado de los objetos mediante una serie de relaciones aplicadas con la definición de Fourier, en el cual cada elemento filtrado de relación es enviado a un sumario de aprendizaje en línea (Danelljan, 2015).

Fase 3: Se generar pequeños videos que son convertidos en un vector vectores de características y se aplica el algoritmo correspondiente.

De esta manera podemos detectar y parametrizar las acciones a realizar en cada uno de los eventos que resulten necesarios con relación al caso de la videovigilancia.

Como explica el autor de la publicación Terence Shin, las ecuaciones están referenciadas al cómputo de las métricas que evalúan y analizan el desempeño del algoritmo, (Shin, 2020).

- TP (True Positive) son los datos Verdaderos positivos
- TN (True Negative) son los datos Verdaderos negativos
- FP (False Positive) son los datos falsos positivos
- FN (False Negative) son los datos falsos negativos
- P son los datos Positivos
- N son los datos negativos
- Precisión = $\frac{VP}{VP, FP, FN}$
- Predicción = datos posibles Positivos y Negativos (P ó N)
- Conjunto vinculado total de los datos prueba = $P+N$

**Figura 5.-** Matriz de Confusión (Shin, 2020)**Ecuaciones de Exactitud, Precisión y Llamado**

Ecuación	
Fórmula de Exactitud	$(T.P. + T.N.) / (P + N)$
Fórmula de Precisión	$T.P. / (T.P. + F.P.)$
Fórmula de Llamado	$T.P. / (T.P. + F.N.)$

Tabla 1.- Definiciones de Fórmulas Matemáticas

En los datos de la tabla 2 se detallan en porcentaje los resultados alcanzados en los indicadores y en la tabla 3 los tiempos de procesamiento del algoritmo.

Medición de Indicadores

Indicador	Valor
Exactitud	80%
Precisión	66%
Recuperación	85%

Tabla 2.- Medición de Indicadores**Tiempo de Procesamiento**

Tiempo del Video (Segundos)	Tiempo de Procesamiento (Segundos)
2	2.25

Tabla 3.- Tiempo de Procesamiento

4. Ventajas y desventajas para la visión por computadora

Debemos considerar que la tecnología maneja ciertas limitantes lo cual se consideran barreras en la aplicación de visión por computador. Entre estas tenemos el registro de su velocidad resultante de la interpretación y la posibilidad que sea susceptible a problemas de relacionados a la iluminación. Adicional a esto debemos tomar en cuenta los costos tanto de mantenimiento como reparación de equipos los cuales suelen ser elevados.

Este tipo de proyectos requiere de tecnología avanzada y precisa por lo tanto son implementados con hardware de alta calidad para la captura de imágenes de alta definición y alto procesamiento de los datos tomando en cuenta que se requiere de ejecución en tiempo real por lo tanto el manejo de servidores en la nube es una excelente opción para cumplir con el hardware requerido. Con la llegada de la tecnología Cloud se podría crear una alta disponibilidad del servicio y así poder consumir bajo demanda de acuerdo con el procesamiento requerido en un lapso determinado.

De igual forma los algoritmos abarcan un alto porcentaje de precisión y mejoran continuamente en un corto tiempo, la única limitante es el tiempo de procesamiento, el cual si es mejorado claramente se denotará en los resultados.

En base a lo antes expuesto la visión artificial contiene métodos tanto de hardware como software que permiten procesar y analizar imágenes generando información en tiempo real encontrado un beneficio tangible al momento de tomar decisiones. Por medio de la tecnología se intenta imitar las funcionalidades de comportamiento de los ojos y cerebro, y mediante el uso de algoritmos poder replicar las decisiones humanas.

Cabe mencionar que como indica el autor, en comparación con la visión que poseemos los humanos, el costo relacionado a la operación de la visión con inteligencia artificial es considerado alto sobre todo cuando es aplicado para bajas cantidades de elaboración (Ramírez Ortiz, 2006).

Según indica Ramírez en su artículo, entre los más destacados contrastes que se presentan de la visión de los humanos y la visión artificial por computadora, obtenemos el modelo de reconocimiento, procesamiento y filtrado de imágenes, orientado hacia la perspectiva de la visión y el proceso de reconocimiento y análisis de los colores. (Ramírez Ortiz, 2006).

En la tabla propuesta que antecede vamos a presentar la comparación de los dos tipos señalados de visión con sus distintas características.

CUADRO COMPARATIVO ENTRE LA VISIÓN HUMANA Y VISIÓN POR COMPUTADORA

Características	Visión Humana	Visión por computadora
Procesamiento en paralelo	Si	No
Procesamiento Serial	No	Si
Perspectiva Tridimensional	Si	No
Perspectiva Bidimensional	No	Si
Fácil reconocimiento de color	Si	No

Tabla 4.- cuadro comparativo evaluando la visión de los humanos y la visión artificial por computadora (Ramírez Ortiz, 2006).

Si bien es cierto que esta tecnología va de la mano con un alto costo económico en infraestructura tecnológica, también aporta como ventaja el desarrollo de tareas repetitivas a grandes velocidades y sin interrupción. Al ser procesos desarrollados por computadoras se reduce el margen de error que suele ocurrir por factores como el cansancio visual o corporal, así como cualquier tipo de distracción que pueda afectar en la productividad o generar errores humanos involuntarios.

Además, visto desde otra perspectiva a largo plazo ayuda en un ahorro económico al requerir de menos personal para la ejecución de procesos, adicional a esto los sistemas trabajan de forma continua en un amplio horario laboral y no es necesario pagar horas extras.

Debemos tomar en cuenta la cantidad de datos de gran relevancia que se pueden recopilar al integrar este tipo de soluciones en sistemas como los puntos de venta donde podemos recopilar información que humanamente sería imposible, como la tasa de conversión de clientes, tomando en cuenta el ingreso de clientes y el porcentaje de clientes que ejecutan las compras, horarios de mayor atención para toma de decisiones estratégicas en cuanto a optimización del número de empleados en horas pico.

Estos datos se obtienen mediante el procesamiento de secuencias de imágenes, analizadas desde diferentes vistas y desde varias cámaras. Esta Visión por computadora artificial se ha incrementado en el uso de propósitos de aplicaciones industriales, en entornos basados en seguridad para la lectura de matrículas de vehículos, manejo de espacios disponibles y el control de tráfico, etc.

Según Solución Ingenieril en su página web indica las siguientes ventajas y desventajas tanto para la visión humana como para la visión artificial por computadora.

VENTAJAS Y DESVENTAJAS DE LA VISIÓN HUMANA

Ventajas de la visión humana	Desventajas de la visión humana
• Radica del conocimiento adquirido a lo largo de la vida, lo cual le brinda información de como es un objeto en cuanto a su tamaño, forma y color y si este es presentado de diferentes formas igual se puede determinar lo que es en base al conocimiento previo. • El ojo humano maneja una amplia resolución de captura de imágenes. • Se adapta con facilidad a las diferentes situaciones como cambien en la iluminación o disposición del escenario. • Facilidad para el reconocimiento de los objetos.	• No cuenta con la habilidad para determinar la magnitud de intensidad de colores • El ojo humano se agota en función del uso realizado, hasta llegar a perder parte de la visión • Al no ser una máquina, el ojo humano es sensible a los efectos ópticos. • No tiene la facultad de identificar distancias geométricas de manera exacta.

Ventajas de la visión artificial	Desventajas de la visión artificial
• Como es un sistema, no se ve influenciado por efectos ópticos. • Tiene la capacidad de medir distancias físicas de forma precisa. • Puede procesar con un alto nivel computacional un gran volumen de imágenes y así obtener información relevante. • Al ser un sistema y no un ser humano, le es posible realizar procesos de forma rutinaria por tiempo prolongado, sin fatigarse y minimizando errores.	• Depende del nivel tecnológico lo cual implica tanto las cámaras utilizadas como el sistema de procesamiento de las imágenes lo cual, la posibilidad de pérdida de información • Los cambios de escala pueden causar cierta dificultad al momento de medir objetos. • Es necesario programar varios escenarios que contemplen los cambios de iluminación para que la interpretación no se vea afectada. • Este tipo de sistemas no logran detectar cuando se realiza el movimiento directamente

Tabla 5.- Ventajas y Desventajas de la Visión Humana vs. Artificial (Ingenieril, 2017)

5. Resultados

Se generaron 215 videos de los cuales solo 80% fueron clasificados de manera correcta. El llamado nos refleja que existe un **85%** de detectar correctamente los eventos.

- En base a los resultados de los datos alcanzados, se puede demostrar la factibilidad de detectar eventos utilizando visión artificial, iniciando con información obtenida a partir de imágenes y videos en situaciones no inspeccionadas (sin controles), por motivo que generalmente provienen de las cámaras CCTV de video vigilancia. Para esto es necesario que las cámaras cuenten con una infraestructura de red que cumpla ciertas características debido a que el procesamiento de la información debe ser en tiempo real.
- Aunque se hayan parametrizado algunas condiciones no controladas, se hace necesario el aplicar algoritmos que lleven a cabo esta medición para que a partir de los mismos se puedan parametrizar los algoritmos de preprocesamiento.
- Como lo sustenta en su investigación la autora Laura Riveros, indica que en los modelos de visión artificial por computadora, la apropiada selección de particularidades que incorporen el objeto analizado o la acción relacionada a detectarse mejora de la misma manera de forma notable su grado de precisión (Riveros, 2019).
- Coincidiendo con la autora Riveros y también con los resultados de una anterior investigación de los autores, se concuerda con lo indicado, que basado en la implantación de soluciones de estas características, se pretende exponer que la utilización de la moderna tecnología actual siendo aplicada de manera correcta es capaz de ayudar significativamente a mejorar los niveles de los controles basados en seguridad y agilizar el trabajo del personal (Cisneros & Quinde, 2019).

6. Conclusiones

En la actualidad se han elaborado varias investigaciones en torno a la visión artificial, inteligencia artificial y la aplicación de esta en el reconocimiento tanto del movimiento como imágenes y en la biometría en general.

Podemos concluir la importancia que existe al aplicar la visión artificial en los sistemas de video vigilancia con reconocimiento facial para así controlar el acceso al personal autorizado. Para lo cual se toma en cuenta los siguientes criterios:

- Eliminación de escenarios con dificultad en relación con la imagen a examinar.
- Seguimiento del movimiento. Las plataformas de software orientadas a entornos de seguridad motivan al desarrollo de habilidades tecnificadas para ejecutar el alcance de seguimiento en entornos controlados y poco o nada controlados.
- Reconstrucción de los datos en movimiento partiendo de variadas vistas.
- Reconstrucción de los datos de movimiento iniciando en imágenes monoculares.
- Reconocimiento de las imágenes obtenidas.

Coincidiendo con lo que indica el autor John Sanabria, se manifiesta que los desarrollos encaminados hacia el análisis de los movimientos humanos a través de la visión por computadora artificial persiguen dar solución una gran cantidad notable de problemáticas proponiendo convertirse en las herramientas tecnológicas de usabilidad para propósito de los diversos estudios, consiguiendo el propósito anhelado (Sanabria, 2011).

Como indica el autor, los trabajos que están orientados a detectar las partes del cuerpo humano deben considerar y tener un funcionamiento y comportamiento invariable, al punto de vista, forma del cuerpo señalado y las condiciones del entorno bajo el cual se está desarrollando el movimiento identificado, es por esto que tienen un gran número de aplicaciones en la ingeniería médica y biomédica, consiguiendo ser motor de los permanentes desarrollos en las áreas orientadas a la captura, análisis e identificación de pequeños y grandes volúmenes de información a través de la visión artificial por computadora (Sanabria, 2011).

El invertir en este tipo de plataformas, desarrollo e investigación de sistemas, permite tener un control preciso de los datos y manejar los accesos en base a un sistema de control. Lo cual minimiza errores y produce información importante para tomar decisiones, adicional a esto, genera un importante ahorro de tiempo al manejar una ejecución continua y efectiva, sin necesidad de pagos de horas extras al personal de seguridad física, ni fallas causadas por el cansancio humano.

Referencias

- Cisneros, E., & Quinde, D. (2019 de Diciembre de 2019). *Revista Carácter*. doi:<https://doi.org/10.35936/caracter.v7i1.57>
- Contactcentersonline. (2019). *contactcentersonline.com*. Obtenido de <https://contactcentersonline.com/inteligencia-artificial-en-seguridad-como-funciona/>

- Culjak, I., & Abram, D. (2013). A brief introduction to OpenCV. (E. a. Everything, Ed.) *MIPRO*, 1725-1730.
- Danelljan, M. (2015). Convolutional features for correlation filter based visual tracking. *In Proceedings of the IEEE international conference on computer vision workshops*, (págs. 58-66).
- Foresti, I. G. (2000). Multimedia video based surveillance systems. *Forensic video investigation Kluwer Academic Publishers*, chapter 1, pages 3–12.
- García Chang, M. E. (2009). *Diseño e implementación de una herramienta de detección facial*. México: Instituto Politécnico Nacional. Obtenido de <http://tesis.ipn.mx/handle/123456789/25106>
- Gebhart, A. (2019). Reconocimiento facial: Apple, Amazon, Google y la carrera por captar tu cara. *CNET*. Obtenido de <https://www.cnet.com/es/noticias/reconocimiento-facial-apple-amazon-google-ai/#:~:text=El%20reconocimiento%20facial%20es%20una,estructura%20%C3%BAnicas%20de%20sus%20rostros>
- Gershgorn, D. (10 de septiembre de 2017). *Quartz*. Obtenido de <https://qz.com/1046350/the-quartz-guide-to-artificial-intelligence-what-is-it-why-is-it-important-and-should-we-be-afraid>
- Gil Jiménez, P. (2009). Estudio, diseño y optimización de técnicas de visión artificial para su aplicación a los sistemas de videovigilancia. Obtenido de <http://hdl.handle.net/10017/6597>
- Gonzalez, R. (2004). *Digital Image Processing using Matlab*. Prentice Hall.
- Ingenieril, S. (2017). *Solucion Ingenieril*. Obtenido de http://solucioningenieril.com/vision_artificial/conceptos_basicos_de_vision_artificial
- Innovación, R. (6 de Junio de 2019). (R. Innovación, Ed.) Obtenido de https://revistainnovacion.com/nota/10498/inteligencia_artificial_en_seguridad_como_funciona/
https://revistainnovacion.com/nota/10498/inteligencia_artificial_en_seguridad_como_funciona/
- Jiménez, P. G. (2009). Estudio, diseño y optimización de técnicas de visión artificial para su aplicación a los sistemas de videovigilancia. *Biblioteca Digital de La Universidad de Alcalá (España)*, <http://hdl.handle.net/10017/6597>. Obtenido de En la Universidad de Alcalá (España) en 2009.
- Muñoz, X. (2003). Image segmentation integrating colour, texture and boundary information. *Universitat de Girona*.
- Ramírez Ortiz, J. (2006). *DESARROLLO DE UNA APLICACIÓN DE INSPECCIÓN*. Universidad de San Carlos de Guatemala, Guatemala. Obtenido de http://biblioteca.usac.edu.gt/tesis/08/08_0313_CS.pdf
- Riveros, L. (2019). *Modelo de visión artificial para la detección automática de colisión de vehículos en videovigilancia*. UNIVERSIDAD NACIONAL DE SAN AGUSTIN DE AREQUIPA.
- Rouhiainen, L. (2018). *Inteligencia artificial*. Madrid: Alienta Editorial.
- Sanabria, J. (2011). Detección y análisis de movimiento usando visión artificial. *Scientia Et Technica*, XVI(49), 180-188. Obtenido de <https://www.redalyc.org/articulo.oa?id=849/84922625031>