

Machine Vision Model Based on Deep Learning to Prevent Theft in Real Estate

Christian Ovalle¹ 

Universidad Tecnológica del Perú, Perú, dovalle@utp.edu.pe

Abstract– Artificial intelligence has become an increasingly present concept today and terms such as Machine Learning or Deep Learning are increasingly used frequently in the development of technological projects and can have an impact on various social problems that, Unfortunately, they are lived day by day in many countries at the Latin American level, which is why the objective of this research was to develop an artificial vision model based on Deep Learning to prevent theft of real estate assets. Likewise, the research is of an applied type and of a quantitative approach, having a pre-experimental cross-sectional design. Finally, after processing the data, it was obtained as results that the convolutional neural networks of Deep Learning are capable of distinguishing the curved lines of a face with very exact precision, only having a margin of error of less than 0.5% in the tests carried out. In addition, it is concluded that the proposed artificial vision model obtained a score in the range of 0.94 to 0.97 for both precision, accuracy, memory and F1, thus proving a positive measurement after applying confusion matrix metrics.

Keywords- Deep Learning, Machine vision, Artificial intelligence, Facial recognition, Neural networks

Modelo de Visión Artificial Basada en Deep Learning para Prevenir Robos en el Patrimonio Inmobiliario

Christian Ovalle¹ 

Universidad Tecnológica del Perú, Perú, dovalle@utp.edu.pe

Resumen– La inteligencia artificial se ha convertido en un concepto cada vez más presente en nuestra actualidad y términos como el Machine Learning o Deep Learning son cada vez más utilizados frecuentemente en el desarrollo de proyectos tecnológicos y pueden generar un impacto en diversas problemáticas sociales que, lamentablemente, se viven día a día en muchos países a nivel de Latinoamérica es por ello, que la presente investigación se planteó como objetivo desarrollar un modelo de visión artificial basada en Deep Learning para prevenir robos en el patrimonio inmobiliario. Asimismo, la investigación es de tipo aplicada y de enfoque cuantitativo, teniendo un diseño preexperimental de corte transversal. Finalmente, luego del procesamiento de los datos se obtuvo como resultados, que las redes neuronales convolucionales del Deep Learning son capaces de distinguir las líneas curvas de un rostro con una precisión muy exacta, solo contando con un margen de error de menos de 0.5% en las pruebas realizadas. Además, se concluye que el modelo de visión artificial propuesto obtuvo un puntaje en el rango de 0.94 a 0.97 tanto para la precisión, la exactitud, la memoria y el F1, comprobándose así una medición positiva luego de aplicar métricas de la matriz de confusión.

Palabras clave- Deep Learning, Visión artificial, Inteligencia artificial, Reconocimiento facial, Redes neuronales

I. INTRODUCCIÓN

Durante los últimos años, la inteligencia artificial se ha convertido en un concepto cada vez más universal en el mundo de la tecnología moderna. En la actual era digital, la inteligencia está llamada a protagonizar Buena parte de los actuales y próximos productos tecnológicos concebidos. Asimismo, en el artículo de Negro & Pons [1] mencionan que la inteligencia artificial es una disciplina próspera que aporta diversos temas de investigación que serán de gran importancia para la resolución de problemas que son intelectualmente complejas para los seres humanos, pero no para las computadoras. Y es que, cuando se habla de inteligencia artificial, se hace referencia, en términos simples, a máquinas que imitan el intelecto del humano para efectuar tareas y con la capacidad de mejorar iterativamente según la información que recopilan. Según Oracle Corporation, una de las más grandes compañías especializadas en el Desarrollo de soluciones

informáticas, tanto locales como en la nube, la inteligencia artificial se ha convertido en un activo empresarial muy valioso gracias a su capacidad de pensamiento y análisis de datos, lo cual ha mejorado el rendimiento y productividad de la empresa mediante la automatización de procesos que antes requerían mayor esfuerzo por parte de las personas.

Cabe mencionar que, los autores Mishra & Pandya [2] en su artículo acerca de la tecnología de las cosas (IoT) se menciona que las expectativas de las tecnologías actuales van en aumento gracias, en parte, a la inteligencia artificial y su relación con técnicas basadas en Machine Learning y Deep Learning. Al respecto, se considera a Machine Learning como un subcampo de la inteligencia artificial y, a su vez, a Deep Learning como un subcampo del Machine Learning. En otras palabras, el Deep Learning es un campo de la inteligencia artificial que ha conseguido favorecer enormemente las famosas redes neuronales con el conocimiento previo del Machine Learning [3].

La cualidad particular del Deep Learning es la capacidad de asemejarse a la forma de pensar de la mente humana para aprender algún tema y mejorar con respecto a los datos que procesa mediante el aprendizaje de ejemplos pre-aprendidos [4]. Gracias a dicha habilidad, el Deep Learning intenta extraer características específicas automáticamente a través de la cantidad de capas secuenciales en su estructura para la creación de un modelo de toma de decisiones en a la resolución de un problema. Este proceso se logra mediante la construcción de modelos computacionales llamados redes neuronales que se inspiran en la estructura del cerebro humano [5]. Por lo tanto, el Deep Learning posee la capacidad de generar nuevos datos después de depender de los datos de entrada almacenados, dotándolo de una gran variedad de posibilidades al momento de la creación de modelos basados en Deep Learning que abarquen diversos problemas o medios de entretenimiento que existen en la sociedad actual.

La tecnología de las redes neuronales del Deep Learning se ha convertido en una de las herramientas más populares de los últimos años. Por ejemplo, las redes generadoras de conflictos, que son los modelos de Deep Learning más conocidas actualmente, pueden producir o redimensionar imágenes en alta calidad, mejorar la calidad de la imagen, convertir imágenes a texto y usarse en seguridad para similares ataques [4]. Cabe mencionar que la inteligencia artificial aplicada a la obtención, procesamiento y análisis de cualquier tipo de información

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).
DO NOT REMOVE

obtenida a través de imágenes digitales, se le conoce como Visión Artificial [6]. En consecuencia, este tipo de inteligencia artificial ha aportado un gran valor para responder de manera más natural a las necesidades y deseos de las personas.

A pesar del crecimiento constante de este tipo de inteligencia artificial y de la ciencia de datos en general, una de las mayores necesidades de los ciudadanos a resolver sigue siendo la inseguridad ciudadana, más específicamente, en países menos desarrollados en donde estas tecnologías, aún, no se han aprovechado de la manera más eficiente posible, ya sea por falta de recursos o desinterés por parte de los gobiernos en financiar proyectos aplicando estos novedosos tipos de algoritmos inteligentes.

En el Perú, no es ninguna novedad de alta tasa de delincuencia que posee el país. Según el Instituto Nacional de Estadística e Informática (INEI), en el periodo de tiempo entre septiembre del año 2021 hasta febrero del 2022, el 86% de la población peruana posee una percepción de inseguridad ciudadana durante los próximos doce meses. Asimismo, a nivel de ciudades de 20 mil a más habitantes en el Perú, sobre todo en la capital Lima, uno de los casos más presente de delincuencia es el robo a los patrimonios, en el que se ha comprobado que casi el 50% de las personas víctimas del delito no denuncian el robo o intento de robo en sus propiedades, siendo una de las principales razones el no poseer evidencia del delincuente.

Con los conocimientos anteriormente mencionados acerca de la inteligencia artificial, es más que un hecho, que los algoritmos inteligentes de la Visión Artificial junto a la tecnología Deep Learning pueden potenciar la creación de modelos y/o sistemas aplicados en los patrimonios inmobiliarios. Como consecuencia, se reduciría la inmensa inseguridad vivida en las residencias en el Perú identificando de formas más eficientes a los delincuentes mediante la entrada de imágenes proporcionadas.

Por lo tanto, es importante realizar un estudio para poder identificar todas las posibilidades beneficiosas que abre la Visión Artificial y el Deep Learning en su contribución para la lucha contra la inseguridad ciudadana en el patrimonio inmobiliario peruano y así responder a la siguiente pregunta ¿Cómo prevenir el alto índice de robos en el patrimonio inmobiliario? Por ello, esta investigación tendrá como finalidad elaborar un modelo de Visión Artificial basado en Deep Learning que pueda ser aplicada en la seguridad del patrimonio inmobiliario.

II. ESTADO DEL ARTE

Se podría decir que la inteligencia artificial (IA) hace mucho tiempo dejó de ser el espectro de la ciencia ficción, para convertirse en la disciplina científica con mayor potencial en los últimos años, del cual existen ciertos campos y aplicaciones que es conveniente resaltar para el Desarrollo de la presente investigación.

A. Visión artificial

La visión artificial como disciplina importante de la IA, nos permite adquirir, procesar y analizar imágenes con el fin de producir información, esta a su vez fue utilizada principalmente en el sector industrial, como fábricas, almacenes, instalaciones, depósitos, que se dedicaban originalmente a la inspección de

etiquetas, la impresión de fecha de producción, la impresión de fecha de caducidad, el control del nivel de líquido de alguna bebida, entre otros. Sin embargo, a la fecha su aplicación no solo está enfocado al sector industrial sino también a diferentes sectores.

Por consiguiente, estudios recientes se han centrado en la visión artificial en humanos, utilizando sus rasgos faciales, gestos y lenguaje corporal [7]. Sin embargo, Obaida, Jamil & Hassan consideran que el Desarrollo de un sistema de visión artificial en humanos que logre la misma efectividad que en el sector industrial sigue siendo difícil [8]. Ya que, autores coinciden que la iluminación, la detección de puntos de referencia faciales o la clasificación son las principales dificultades en este entorno [8] [9].

Para finalizar, el Desarrollo de los diferentes sistemas inteligentes enfocado en la visión artificial, se ha convertido en una de las direcciones de innovación más importantes en el campo de los chips de la inteligencia artificial [10].

B. Deep Learning

El Deep Learning (DL) también conocido como aprendizaje profundo, es considerado como un tipo de machine learning, que hace uso de las redes neuronales profundas para entrenar a un computador, el cual está emergiendo en los últimos años como una poderosa herramienta para el aprendizaje automático, que promete remodelar el futuro de la IA, haciendo que el algoritmo aprenda una gran muestra de datos para representar la variable de respuesta con éxito [11] [12].

Por su lado, diversos autores denotan que los modelos en base a Deep learning, son capaces de realizar una predicción correcta, a partir de datos de entrada y gracias a esto es posible entrenar tales modelos con el fin de que puedan reconocer un único objetivo arbitrario [13].

Por el contrario, aunque existen diferentes tareas de IA, aplicando las técnicas Deep Learning que pueden ofrecer mejoras sustanciales, en comparación con los enfoques tradicionales de los automáticos, muchos profesionales del sector salud se muestran escépticos en cuanto a su uso, esto se debe a que las bases del aprendizaje profundo aún no han proporcionado soluciones completas [13].

C. Seguridad en el hogar inteligente

Con el desarrollo de nuevos avances tecnológicos, tal como lo son la visión artificial y las técnicas de Deep Learning, surge la necesidad de un sistema que refuerce la seguridad en nuestras propiedades inmobiliarias, ya que hoy en día, contar con un sistema de seguridad es considerado un tema crucial, el cual garantiza que las propiedades y los seres queridos estén siempre seguros y protegidos [14].

En Malasia, un grupo de estudiantes consideraron que un sistema de seguridad se ha vuelto vital para prevenir y garantizar la seguridad en los hogares, por lo que propusieron para superar este tema preocupante, la implementación de un sistema de reconocimiento de rostros aplicando aprendizaje

profundo e internet of thing (IoT), que permitiera al usuario el poder controlar el acceso de su puerta mediante la visión artificial [14].

En la misma línea, muchos investigadores recomiendan el uso de un dispositivo llamado Raspberry PI, para poder implementar este tipo de sistema de visión artificial, ya que consideran que dicho dispositivo tiene un alto límite de manejo, un bajo costo y se ajusta a varios modos de programación [15].

Por otro lado, en una Universidad de Indonesia, un grupo de investigadores, desarrollaron un sistema de detección e identificación humana para el sistema de monitoreo del hogar, teniendo como base el método de histograma de gradiente, que es capaz de identificar si un objeto es un humano o no. Según, las pruebas realizadas en objetos no humanos, como animales y objetos que pudiera asemejar a los humanos, el sistema de monitoreo hacienda uso de la visión artificial pudo identificar que efectivamente no se trataba de un ser humano [16].

Finalmente, el uso de un sistema de monitoreo de visión artificial para prevenir robos en el patrimonio inmobiliario, es algo que muchas familias deberían implementar en sus propiedades, ya que nos puede brindar muchas ventajas y automatizar procesos. Sin embargo, en la actualidad esta tecnología tiene un costo muy alto por lo que muchas personas prefieren omitir este gasto.

III. METODOLOGÍA

El presente artículo de investigación abarca una investigación aplicada de enfoque cuantitativo de diseño preexperimental de corte transversal, para Sánchez [17] la investigación bajo el enfoque cuantitativo trata con fenómenos que se pueden medir, a través de diversas técnicas estadísticas para el análisis de los datos recogidos y para su posterior respuesta a la pregunta de investigación. Por su lado, Otero [18] señala que el enfoque cuantitativo se basa en experimentos y análisis de causa y efectos. Es decir, se realizará de manera experimental con técnicas de diseño de observación, ya que el modelo de visión artificial propuesto involucre investigación en el aspecto tecnológico y su aplicación de este en la prevención de robos en el patrimonio inmobiliario. Asimismo, Lozada [19] menciona a la investigación aplicada como un proceso que permite transformar el concepto científico o empírico en conceptos, prototipos y/o productos.

Con el fin de desarrollar el modelo de visión artificial basada en Deep Learning para prevenir robos en el patrimonio inmobiliario, se realizó el diseño del modelo que es definida por los procesos a ejecutar al momento de la utilización del modelo habiendo sido implementado en un sistema de visión artificial cumpliendo, finalmente, el objetivo principal del tema de investigación. En primer lugar, este modelo podrá ser implementado en cualquier sistema que permita monitorear en tiempo real y funcionar las 24 horas del día, con la finalidad de vigilar la propiedad inmobiliaria captando imágenes de posibles sospechosos. Las imágenes serán enviadas al sistema programado con la inteligencia artificial del Deep Learning, más específicamente con los algoritmos tipo redes neuronales. Este tipo de algoritmos de Deep Learning están diseñadas para trabajar con imágenes en el proceso de clasificación y segmentación de estas [20]. Por ende, las redes neuronales son

utilizadas en los modelos de visión artificial de forma recurrente en los proyectos realizados en los últimos años [20]. Estos algoritmos tomarán de datos de entrada de entrenamiento a las imágenes captada por el dispositivo de grabación y darán inicio al análisis de las capturas obtenidas para el reconocimiento de un posible intruso. Una vez reconocido el sospechoso, la imagen será segmentada por los algoritmos validando la presencia del intruso y, finalmente, la imagen será enviada al dispositivo de monitoreo del usuario alertando de la situación de riesgo. Todo el procedimiento se muestra en la fig. 1.

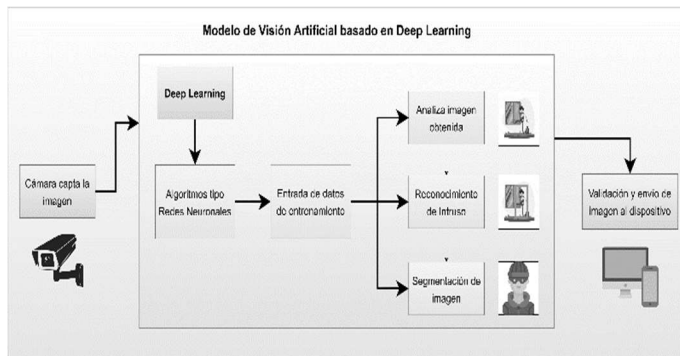


Fig. 1 Diseño del modelo de Visión Artificial basado en Deep Learning

Las redes neuronales convolucionales (CNN), para identificar y clasificar objetos en imágenes, posee varias capas especializadas en la detección de líneas, curvas, etc, que existen en estas, y poco a poco van especializándose hasta llegar a capas más profundas que reconocen formas complejas como un rostro humano. Las redes neuronales, como se mencionó anteriormente, trabajan con capas que analizan los datos recibidos en el modelo, en específico, son tres tipos de capas las que existen: La capa de entrada, recibe los datos de entrada y los pasa a la primera capa oculta; Las capas ocultas, realizan los cálculos programados con los datos de entrada; La capa de salida, devuelve la predicción realizada luego de que el proceso de las capas ocultas haya concluido. Para el caso de nuestro modelo, se utilizará la librería de código “Keras”, que proporciona la capacidad de crear bloques modulares que contienen las redes neuronales con sus capas, sobre las que se pueden desarrollar modelos complejos de aprendizaje profundo [20].

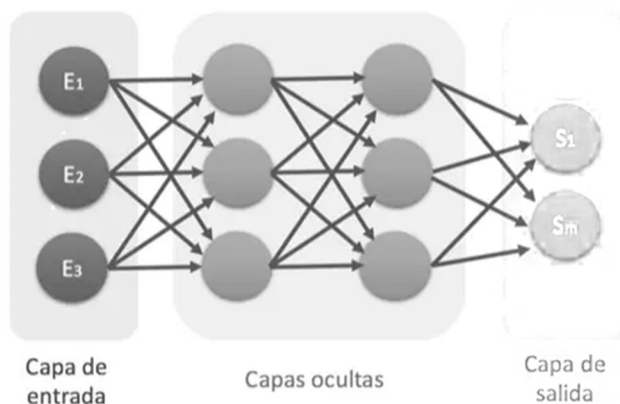


Fig. 2 Capas de un algoritmo tipo Red Neuronal

Los algoritmos tipo redes neuronales, como toda inteligencia artificial, necesita de un entrenamiento a base de una gran cantidad de datos para un mejor funcionamiento. En el caso de nuestro modelo, se crearán bases de datos que almacenarán muchas imágenes que servirán como datos de entrenamiento para que los algoritmos procedan a empezar a diferenciar a las personas que pueden ser posibles sospechosos de los propietarios del patrimonio inmobiliario.

Para el desarrollo de dicho modelo de visión artificial se ha considerado la estructura de trabajo expuesta en la Figura 1, que expone los distintos procesos del modelo de visión artificial. Asimismo, en las siguientes líneas se detallarán con mayor amplitud cada una de las fases que se llevaron a cabo en el presente trabajo de investigación:

Fase 1: Herramientas a utilizar para la implementación del modelo de visión artificial:

- Se eligió el lenguaje de programación Python versión 3.10.5.
- Se utilizó el entorno de Desarrollo integrado PyCharm, el cual es utilizado concretamente para el lenguaje de programación Python.
- Una cámara web para capturar las imágenes que servirán de datos de entrada para los algoritmos.
- Las librerías a instalar son parte fundamental para la implementación del modelo, entre las más destacadas están las librerías de la biblioteca OpenCV, la cual proporciona herramientas dedicadas a la visión artificial; Numpy, que nos permite hacer operaciones matriciales; Imutils, nos va permitir procesar imágenes; Keras, librería que nos va a permitir ir construyendo cada bloque de cada red neuronal; Scikit-learn, librería que posee herramientas para la medición del desempeño de los algoritmos del Deep Learning; entre otras librerías.

Fase 2: Base de datos de Imágenes

Para el presente modelo de visión artificial se crearán 2 bases de datos, la primera almacenara las fotografías de los usuarios que puedan habitar sea una casa, un local o un negocio, el cual se le denominara como "Usuarios". Seguidamente, viene la base de datos denominada como "Desconocidos" que almacenara fotografías de diferentes rostros. En las siguientes líneas se detalla lo antes expuesto:

Para la base de datos Usuarios, se creó un algoritmo capaz de realizar tomas masivas de fotografías a través de la cámara web, esto con el fin de obtener una base de datos solida que identifique y guarde a los posibles usuarios del patrimonio inmobiliario y que posteriormente el modelo no lo identifique como intruso. Asimismo, este algoritmo también nos ayudara a que las imágenes se encuentren debidamente etiquetadas. Es así que se logró recaudar un total 667 fotografías al rostro de los usuarios, con diferentes posiciones.



Fig. 3 Base de datos de Usuario

Para la base de datos llamada Desconocidos, se obtuvieron fotografías de distintos rostros de personas. Estas fotografías fueron descargadas de internet mediante una extensión propia de Google Chrome que permite la descarga masiva de las mismas. Asimismo, haciendo uso del algoritmo se pudieron etiquetar las fotografías obtenidas de internet. Por lo que se lograron recolectar 1610 imágenes de fotografías de distintos rostros de hombres y mujeres.



Fig. 4 Base de datos Desconocido

Fase 3: Construir el Modelo

- Para el diseño del presente modelo de visión artificial se ha considerado una red neuronal convolucional (CNN) que tiene 2 capas convolucionales y 3 capas rectificadas.
- Para comenzar con el proceso de entrenamiento de nuestro modelo se ha de crear una dirección que mapee las fotografías recopiladas tanto para usuarios como para desconocidos.
- Seguidamente llama a la colección de imágenes que se tiene en la base de datos, para comenzar a guardar la cantidad de imágenes que se tiene en las 2 base de datos.
- Posteriormente se guarda el modelo ya entrenado con sus pesos.
- Se realiza el entrenamiento del modelo asignándole 20 épocas. En la figura 4 se pueda visualizar como el modelo está siendo entrenado y a la vez se puede validar la métrica accuracy para cada época que está por encima del 0.9 que es lo que identifica que la precisión está bordeando la unidad que es el valor máximo para medir un modelo.

```

Epoch 1/20
46/46 [.....] - ETA: 0s - loss: 0.1616 - accuracy: 0.9350WARNING:absl:Found untraced functions such as _jit_compil
46/46 [.....] - 15s 315ms/step - loss: 0.1616 - accuracy: 0.9350 - val_loss: 0.0250 - val_accuracy: 0.9891
Epoch 2/20
46/46 [.....] - 14s 389ms/step - loss: 0.0422 - accuracy: 0.9890 - val_loss: 0.0315 - val_accuracy: 0.9863
Epoch 3/20
46/46 [.....] - ETA: 0s - loss: 0.0214 - accuracy: 0.9932WARNING:absl:Found untraced functions such as _jit_compil
46/46 [.....] - 15s 337ms/step - loss: 0.0214 - accuracy: 0.9932 - val_loss: 0.0868 - val_accuracy: 0.9973
Epoch 4/20
46/46 [.....] - ETA: 0s - loss: 0.0109 - accuracy: 0.9973WARNING:absl:Found untraced functions such as _jit_compil
46/46 [.....] - 14s 366ms/step - loss: 0.0109 - accuracy: 0.9973 - val_loss: 0.0929 - val_accuracy: 1.0000
Epoch 5/20
46/46 [.....] - 12s 270ms/step - loss: 0.0254 - accuracy: 0.9904 - val_loss: 0.0842 - val_accuracy: 1.0000
Epoch 6/20
46/46 [.....] - ETA: 0s - loss: 0.0888 - accuracy: 0.9952WARNING:absl:Found untraced functions such as _jit_compil
46/46 [.....] - 13s 288ms/step - loss: 0.0888 - accuracy: 0.9952 - val_loss: 1.4035e-04 - val_accuracy: 1.0000
Epoch 7/20
46/46 [.....] - ETA: 0s - loss: 0.0861 - accuracy: 0.9979WARNING:absl:Found untraced functions such as _jit_compil
46/46 [.....] - 13s 285ms/step - loss: 0.0861 - accuracy: 0.9979 - val_loss: 1.1128e-04 - val_accuracy: 1.0000
Epoch 8/20
46/46 [.....] - 12s 262ms/step - loss: 0.0024 - accuracy: 0.9993 - val_loss: 1.3143e-04 - val_accuracy: 1.0000

```

Fig. 5 Fase de entrenamiento “Épocas”

Fase 4: Evaluar el Modelo

Se evalúa el modelo de visión artificial según los resultados obtenidos y el cumplimiento de los objetivos propuestos. Asimismo, se utiliza la librería Scikit-learn para medir el desempeño de la métrica de precisión de los algoritmos del Deep Learning implementados en el modelo.

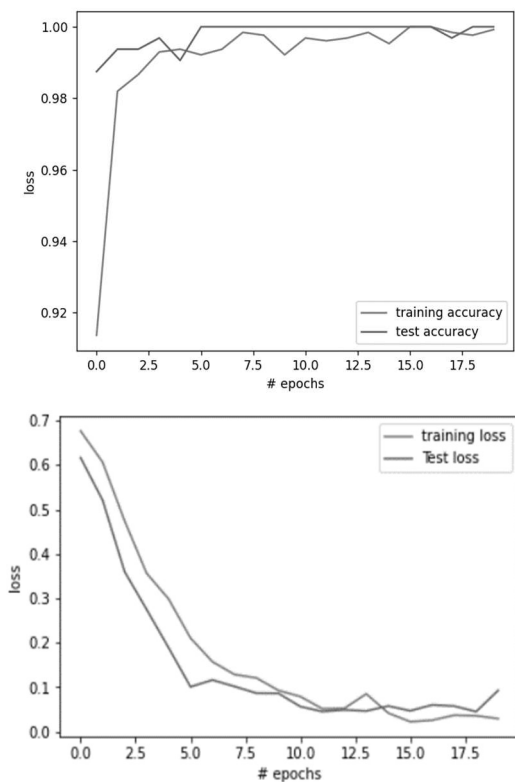


Fig. 6 Evaluación de precisión de un modelo

Fase 5: Implantación

En esta última fase se explica la funcionalidad del modelo construido en las fases anteriores. Asimismo, indicar los resultados obtenidos de forma entendible. Por último, realizar el informe en el que se incluya posibles mejoras para el futuro.

IV. RESULTADOS

Una vez construido y entrenado nuestra propuesta tecnológica de modelo de visión artificial basado en Deep

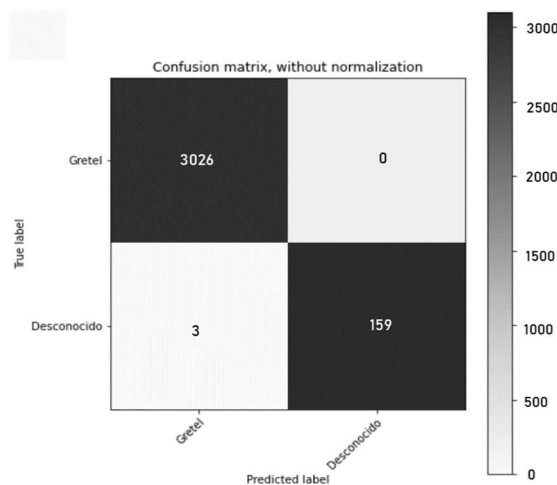
Learning para prevenir robos en el patrimonio inmobiliario, llega el momento de evaluarlo y conocer si nuestro modelo identifica y clasifica correctamente las imágenes. Para el proceso de evaluación del modelo se utilizará la biblioteca scikit-learn, la cual contiene algoritmos de medición que es comúnmente utilizado para evaluar el rendimiento de modelos de Machine Learning y Deep Learning. Scikit-learn permite la realización de una matriz de confusión, la cual es una representación de los valores predictivos que posee el modelo al momento de su funcionamiento.

		Predicted Values	
		Negative	Positive
Actual Values	Negative	TN True Negative	FP False positive
	Positive	FN False Negative	TP True Positive

Fig. 7 Matriz de Confusión

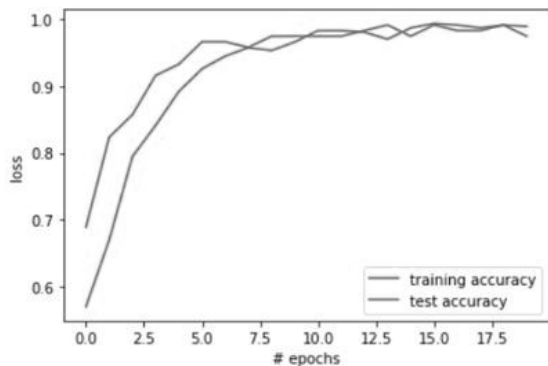
Cada valor de predicción puede ser uno de los cuatro resultados, basado en cómo coincide con el valor real, en el caso de nuestro modelo, la identificación de un rostro conocido y de uno desconocido.

Se realizó una primera prueba del modelo de visión artificial utilizando 3188 imágenes, dando como resultado la matriz de confusión en forma de mapa de calor que se muestra en la figura 7.



Dado el resultado de las 3188 imágenes procesadas, se obtuvo: 3026 imágenes con predicción verdadero e identificado verdadero; 0 imágenes con predicción verdadero e identificado negativo; 3 imágenes con predicción negativo e identificado verdadero; 159 imágenes con predicción negativo e identificado negativo. Estos datos son representados en la

gráfica que se muestra en la figura 8, en el cual se observa la comparación de la tendencia de la precisión del modelo durante el entrenamiento y la precisión al momento de realizar la prueba. Como resultado, se muestra que la precisión general del modelo es eficiente con un valor de 0.96 (96%) al momento de clasificar imágenes de rostros conocidos y desconocidos.



Accuracy Score = 0.9645669291338582
Precision Score = 0.9338513381336513

Fig. 9 Gráfica de precisión del modelo de visión artificial

En base a los resultados de la matriz de confusión, podemos calcular las métricas específicas del modelo de visión artificial. Dichas métricas son: Exactitud, Precisión, Memoria y F1 (precisión en un test). Las fórmulas de cada métrica se muestran en la figura 9.

$$\text{Exactitud} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precisión} = \frac{TP}{TP + FP}$$

$$\text{Memoria} = \frac{TP}{TP + FN}$$

$$F1 = 2 * \frac{\text{Precisión} * \text{Memoria}}{\text{Precisión} + \text{Memoria}}$$

Fig. 10 Fórmulas de métricas

Por último, las métricas calculadas se ingresarán en el instrumento creado para la evaluación del modelo de visión artificial. Los resultados obtenidos son: En exactitud se obtuvo un valor de 0.94; En memoria se obtuvo un valor de 0.95; En precisión se obtuvo un valor de 0.99; En F1 se obtuvo un valor de 0.97.

TABLA I
MÉTRICAS OBTENIDAS EN LA MEDICIÓN DEL MODELO

Nº	Resultados	Matriz y Métricas				Observaciones
		Exactitud	Memoria	Precisión	F1	
1	Medición en base a 3188 imágenes procesadas.	0.94	0.95	0.99	0.97	Se comprobó una medición positiva del funcionamiento

						del modelo.
--	--	--	--	--	--	-------------

TABLA II

MÉTRICAS OBTENIDAS EN LA MEDICIÓN DEL MODELO SEGÚN LAS ÉPOCAS

Iteraciones	Matriz y Métricas			
	Exactitud	Memoria	Precisión	F1
1	0.55	0.58	0.58	0.57
2	0.72	0.73	0.75	0.76
3	0.86	0.88	0.80	0.88
4	0.91	0.93	0.95	0.96
5	0.94	0.95	0.99	0.97

V. DISCUSIÓN

Por consiguiente, es preciso indicar y partir de la pregunta de investigación del ¿Cómo prevenir el alto índice de robos en el patrimonio inmobiliario?, que tiene como finalidad el diseñar un modelo de visión artificial basado en Deep Learning, empleando redes neuronales convolucionales. De esta manera, se propuso un modelo teniendo como base el flujo de la integración del Deep learning, en el reconocimiento de intrusos en el patrimonio inmobiliario.

Por lo cual, para el entrenamiento del modelo en mención, se parte de una gran cantidad de imágenes debidamente recopiladas y que a su vez se acopla a diversas capas y niveles de procesamiento, propias de las redes neuronales convolucionales. Sin embargo, conforme se iba diseñando el modelo se fueron realizando ajustes y mejoras en base al objetivo de la investigación.

Asimismo, es indispensable mencionar que mientras se realizaban ajustes en el modelo, también se incrementaban la cantidad de imágenes en el dataset, tomando como referencia investigaciones preliminares donde afirmaban que, a mayor cantidad de imágenes procesadas, mayor era la precisión o exactitud de los resultados del modelo.

VI. CONCLUSIONES

Según los resultados obtenidos se concluye que la tecnología de los algoritmos tipo Redes neuronales convolucionales del Deep Learning otorga una gran precisión al momento de identificar, tanto a los dueños de la propiedad inmobiliaria como a posibles sospechosos. Con ayuda de un gran almacenamiento de imágenes para un entrenamiento eficiente, las redes neuronales convolucionales del Deep Learning son capaces de distinguir las líneas curvas de un rostro con una precisión muy exacta, solo contando con un

margen de error de menos de 0.5% en las pruebas realizadas.

De manera análoga, es preciso concluir que el modelo de visión artificial propuesto obtuvo un puntaje en el rango de 0.94 a 0.97 tanto para la precisión, la exactitud, la memoria y el F1, comprobándose así una medición positiva, ya que bordea cerca de la unidad, que es el número máximo que se toma en consideración al medir modelos de inteligencia artificial.

Del mismo modo podemos afirmar que, el uso de la inteligencia artificial para el apoyo contra problemáticas sociales como la inseguridad ciudadana, es necesario para romper las barreras tecnológicas y dejar al lado el rechazo al uso de las nuevas tecnologías para mejorar la calidad de vida de la población.

REFERENCIAS

- [1] P. Negro y C. Pons, «Artificial Intelligence techniques based on the integration of symbolic logic and deep neural networks: A systematic literature review», *Inteligencia Artificial*, vol. 25, n.º 69, pp. 13-41, jun. 2022, doi: 10.4114/intartif.vol25iss69pp13-41.
- [2] N. Mishra y S. Pandya, «Internet of Things Applications, Security Challenges, Attacks, Intrusion Detection, and Future Visions: A Systematic Review», *IEEE Access*, vol. 9, pp. 59353-59377, abr. 2021, doi: 10.1109/ACCESS.2021.3073408.
- [3] A. Caroppo, A. Leone y P. Siciliano, «Comparison Between Deep Learning Models and Traditional Machine Learning Approaches for Facial Expression Recognition in Ageing Adults», *Journal of Computer Science and Technology*, vol. 35, n.º 5, pp. 1127-1146, oct. 2020, doi: 10.1007/s11390-020-9665-4.
- [4] C. Ma, «Smart city and cyber-security; technologies used, leading challenges and future recommendations», *Energy Reports*, vol. 7, pp. 7999-8012, nov. 2021, doi: 10.1016/j.egy.2021.08.124.
- [5] L. Jin, F. Tan y S. Jiang, «Generative Adversarial Network Technologies and Applications in Computer Vision», *Hindawi: Computational Intelligence and Neuroscience*, vol. 2020, n.º 1459107, ago. 2020, doi: 10.1155/2020/1459107.
- [6] S. Vallabhaneni, «Image classification for user feedback using Deep Learning Techniques», *5th International Conference on Computing Methodologies and Communication, ICCMC 2021*, abr. 2021, doi: 10.1109/ICCMC51019.2021.9418419.
- [7] S. S. Farfade, M. Saberian, y L.-J. Li, «Multi-view Face Detection Using Deep Convolutional Neural Networks». arXiv, 20 de abril de 2015. Accedido: 13 de mayo de 2022. [En línea]. Disponible en: <http://arxiv.org/abs/1502.02766>.
- [8] T. H. Obaida, A. S. Jamil, y N. F. Hassan, «Real-time face detection in digital video-based on Viola-Jones supported by convolutional neural networks», *IJECE*, vol. 12, n.º 3, p. 3083, jun. 2022, doi: 10.11591/ijece.v12i3.pp3083-3091.
- [9] X. Zhao y S. Zhang, «Facial Expression Recognition Based on Local Binary Patterns and Kernel Discriminant Isomap», *Sensors*, vol. 11, n.º 10, pp. 9573-9588, oct. 2011, doi: 10.3390/s111009573.
- [10] Q.-B. Zhu et al., «A flexible ultrasensitive optoelectronic sensor array for neuromorphic vision systems», *Nat Commun*, vol. 12, n.º 1, p. 1798, dic. 2021, doi: 10.1038/s41467-021-22047-w.
- [11] Y. Ghasemi, H. Jeong, S. H. Choi, K.-B. Park, y J. Y. Lee, «Deep learning-based object detection in augmented reality: A systematic review», *Computers in Industry*, vol. 139, p. 103661, ago. 2022, doi: 10.1016/j.compind.2022.103661.
- [12] D. Ravi et al., «Deep Learning for Health Informatics», *IEEE J. Biomed. Health Inform.*, vol. 21, n.º 1, pp. 4-21, ene. 2017, doi: 10.1109/JBHI.2016.2636665.
- [13] A. Ferrato, C. Limongelli, M. Mezzini, y G. Sansonetti, «Using Deep Learning for Collecting Data about Museum Visitor Behavior», *Applied Sciences*, vol. 12, n.º 2, p. 533, ene. 2022, doi: 10.3390/app12020533.
- [14] S. A. Radzi, M. K. M. F. Alif, Y. N. Athirah, A. S. Jaafar, A. H. Norihan, y M. S. Saleha, «IoT based facial recognition door access control home security system using raspberry pi», *IJPEDS*, vol. 11, n.º 1, p. 417, mar.

- 2020, doi: 10.11591/ijpeds.v11.i1.pp417-424.
- [15] «Real Time Access Control Based on Face Recognition», presentado en *International Academy of Engineers*, jun. 2015. doi: 10.15242/IAE.IAE0615004.
- [16] D. Arisandi, M. Elveny, y R. Rahayu, «Human Detection and Identification for Home Monitoring System», *J. Phys.: Conf. Ser.*, vol. 1898, n.º 1, p. 012026, jun. 2021, doi: 10.1088/1742-6596/1898/1/012026.
- [17] Sánchez Flores, Fabio Anselmo, «Fundamentos epistémicos de la investigación cualitativa y cuantitativa: consensos y disensos», *Revista Digital de Investigación en Docencia Universitaria*, vol. 13, n.º 1, pp. 102-122, ene. 2019, doi: 10.19083/ridu.2019.644.
- [18] A. Otero Ortega, «ENFOQUES DE INVESTIGACIÓN», p. 35.
- [19] J. O. Lozada, «Investigación Aplicada: Definición, Propiedad Intelectual e Industria», p. 6, 2014.
- [20] J. Schmirhuber, «Deep learning in neural networks: An overview», *Neural Networks*, vol. 61, p. 85-117, ene. 2015, doi: <https://doi.org/10.1016/j.neunet.2014.09.003>.