




Performance of Image Recognition with Machine Vision: A Systematic Review

Joseph Lopez-Carreño, Bach.¹, Cristhian Calvo-Lavado, Bach.¹, and Eliseo Zarate-Perez, PhD¹




¹Facultad de Ingeniería, Universidad Privada del Norte (UPN), Perú, N00165247@upn.pe, N00179564@upn.pe, eliseo.zarate@upn.edu.pe

Abstract– This study aimed to identify the existing techniques, applications, equipment, and technologies applied for recognizing images with artificial vision via a systematic review of the literature during the period 2020–2022. PRISMA was used for selecting and analyzing 142 articles obtained from the EBSCO, Engineering Source, ProQuest, and ScienceDirect databases. Studies that were not directly related to the proposed objectives were not included, leaving 28 articles for full-text review. The review results strongly suggest that Hopfield-type convolutional artificial neural networks are highly effective for image recognition and classification tasks. Similarly, the combination of technological tools such as YOLO, Roboflow, Python, and OpenCV shows that image processing and deep learning are driving new applications that improve the various performance metrics of these tasks. Therefore, artificial vision, unlike technologies that incorporate electronic devices with sensors, allows the interpretation of an environment with a high degree of representation of reality, confirming its robustness in the complexity of data processing.

Keywords– artificial intelligence, computer vision, image recognition, neural networks, object detection.

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

Desempeño del Reconocimiento de Imágenes con Visión artificial: Una Revisión Sistemática

Joseph Lopez-Carreño, Bach.¹, Cristhian Calvo-Lavado, Bach.¹, y Eliseo Zarate-Perez, PhD¹

¹Facultad de Ingeniería, Universidad Privada del Norte (UPN), Perú, N00165247@upn.pe, N00179564@upn.pe, eliseo.zarate@upn.edu.pe

Resumen– El objetivo de este trabajo fue identificar las técnicas, aplicaciones, equipos y tecnologías existentes aplicados en el reconocimiento de imágenes con visión artificial a través de una revisión sistemática de la literatura durante el periodo 2020 – 2022. Para el trabajo de revisión se utilizó la declaración PRISMA en la selección y análisis de los 142 artículos obtenidos de las bases de datos científicas EBSCO, Engineering Source, ProQuest y ScienceDirect. Se descartaron estudios que no tengan relación directa con los objetivos planteados, identificándose 28 artículos para la revisión a texto completo. Los resultados de revisión muestran que las redes neuronales artificiales convolucionales de tipo Hopfield tienen una alta efectividad para realizar tareas de reconocimiento y clasificación de imágenes. De la misma forma, la combinación de herramientas tecnológicas como YOLO, Roboflow, Python y OpenCV muestran que el procesamiento de imágenes y el aprendizaje profundo están impulsando nuevas aplicaciones que mejoran las diversas métricas de desempeño en estas tareas. Por lo tanto, la visión artificial, a diferencia de las tecnologías que incorporan dispositivos electrónicos con sensores, permite una interpretación del entorno con un mayor grado de representación de la realidad, siendo robustos a la complejidad del procesamiento de los datos.

Palabras clave– Inteligencia artificial, visión artificial, reconocimiento de imágenes, redes neuronales, detección de objetos.

I. INTRODUCCIÓN

Alrededor de la década de los 50 se iniciaron las primeras investigaciones de visión por computador. Desde aquel entonces han aparecido diversos métodos a partir de los cuales se han ido mejorando los resultados de procesamiento de imágenes [1]. La visión por computadora (CV) es el campo de estudio que se ocupa de cómo las computadoras pueden analizar imágenes o videos digitales y buscar automatizar tareas que puede realizar el sistema visual humano. Debido a que este campo se ocupa de todos los procesos de adquisición de información del mundo real por parte de las computadoras, la visión artificial es extensa y abarca desde equipos para detección de imágenes hasta reconocimiento de imágenes basado en IA [2].

La inteligencia artificial, en particular las redes neuronales artificiales orientadas al aprendizaje profundo, como el *Deep Learning* o *Machine Learning*, son estructuras computacionales que pretenden igualar las funciones ejecutadas por el cerebro humano. Estas poseen la habilidad de aprender a través de un entrenamiento y se han convertido en un instrumento muy utilizado por su amplia variedad de aplicaciones en muchas áreas de la ciencia y la salud [3]. Las aplicaciones de IA han experimentado un crecimiento sin precedentes en las últimas décadas debido a la mejora del poder computacional y la

disponibilidad de grandes conjuntos de datos [4]. Además, ellos permiten resolver problemas como el análisis de patrones, el reconocimiento de imágenes o la búsqueda de diferencias entre objetos. Sin embargo, la identificación automática de objetos en una imagen sigue siendo un tema emergente en la visión por ordenador [5].

Aunque hay muchas aplicaciones de procesamiento de imágenes, los principios y métodos empleados en ellas son los mismos y no hay diferencia. Detrás de la alta eficiencia y precisión, todavía hay incertidumbre en la tecnología de visión por computador. En la evolución de la tecnología de visión por computador, la incertidumbre inherente de los mecanismos operativos hace que el desarrollo de la tecnología sea similar a un estado de “caja negra”, por lo que es difícil para los investigadores encontrar el error y controlar el rendimiento del producto.

La incertidumbre de las variables exógenas se refiere principalmente al riesgo de desarrollo tecnológico, dentro de los cuales el riesgo de costo y el riesgo de cronograma son los de mayor impacto. En el caso de la inteligencia artificial, incluso si su tecnología puede desarrollarse como se espera, aún se enfrenta a un juicio sobre esta tecnología [6]. Con todo ello, en [7] se refiere a que después de años de investigación, los humanos han realizado exploraciones más profundas en la industria de la tecnología de visión por computador y han cosechado muchos resultados con significado práctico.

Las aplicaciones de visión por computador se han expandido rápidamente desde escenarios de aplicaciones específicas para convertirse en una herramienta científica estándar y aplicarse a aplicaciones de tecnología y ciencias naturales. Ello es debido al desarrollo de las técnicas de inteligencia artificial, la mejora del poder computacional y la disponibilidad de grandes conjuntos de datos [8]. Según [4], en el campo de la medicina se puede usar para el pronóstico y la interpretación de imágenes médicas.

Según [9], en el campo del reconocimiento facial los investigadores pueden mejorar la precisión de un gran número de muestras de imágenes faciales simplemente pasándolas a un modelo de aprendizaje profundo. Entre otras aplicaciones, la visión artificial tiene las ventajas de una alta precisión de reconocimiento, una velocidad de detección rápida y un gran potencial de desarrollo [10]. Ella ha logrado un éxito considerable en la clasificación de imágenes, detección de objetos, segmentación de imagen y rostro. Además, se han utilizado ampliamente en agricultura, atención médica, educación, inspección industrial y otros campos.

La visión artificial es una clasificación de la inteligencia artificial que se define como una agrupación de diferentes métodos, técnicas o algoritmos que permiten el procesamiento, análisis y traducción de cualquier tipo de patrones o información contenida en las imágenes digitales [2]. Existen múltiples técnicas de IA, como Aprendizaje Automático (ML), Redes neuronales artificiales (ANN), Proceso Natural de Lenguaje (NLP), Sistemas Expertos, Visión Artificial (AV), entre otros. De igual manera, algunas técnicas mencionadas agrupan a un subconjunto de algoritmos, en el caso del ML agrupa al algoritmo Deep Learning y los algoritmos Supervisado y No supervisado, en el caso de la ANN tiene como clases a las redes neuronales convolucionales, Dinámico, Binario, Feedforward, Recurrente, de Hopfield e Hídrico, entre otros [4].

Los algoritmos y técnicas mencionados necesitan de modelos que estudien todas las características necesarias de las imágenes para indicar la presencia de objetos, colores, animales o personas, lo que se define a todas ellas como datos de entrenamiento [1]. De esta forma, para el entrenamiento de grandes bases de datos, ha sido necesario el crecimiento de las unidades de procesamiento gráfico (GPU), las cuales han ayudado a disminuir considerablemente los tiempos de entrenamiento de los modelos de red [3]. Además, en [11] se refiere la alta complejidad de la gran cantidad de cálculos y procesos realizados, especialmente durante el entrenamiento de datos aportados a las redes neuronales, obliga a que se requiera componentes físicos. Ello se relaciona con el hardware de grandes capacidades para soportar cargas de trabajo intensivas aprovechando múltiples recursos con base en una eficiencia energética.

Por consiguiente, el presente trabajo pretende responder a la interrogante: ¿Cuáles son los hallazgos acerca de las técnicas, aplicaciones, equipos y tecnologías existentes para lograr una alta funcionalidad de reconocimiento de imágenes procesadas por visión artificial reportados durante el periodo 2020 – 2022?

En este aspecto, el objetivo de este trabajo fue identificar las técnicas, aplicaciones, equipos y tecnologías existentes aplicados en el reconocimiento de imágenes con visión artificial a través de una revisión sistemática de la literatura durante el periodo 2020 – 2022.

La información recolectada en esta investigación será de gran utilidad para todas aquellas personas que quieran aprender acerca de las técnicas, equipos y tecnologías recientes en el campo de la visión por computador, incluyendo las diferentes aplicaciones en la que estas se emplean en la actualidad. Asimismo, servirles de orientación en el empleo pertinente de la técnica deseada, con el fin de poder analizar y detectar patrones contenido en las imágenes [12].

II. METODOLOGÍA

Este trabajo de revisión sistemática se realizó siguiendo la guía práctica del flujograma de la Declaración PRISMA, siguiendo la metodología propuesta en [13]–[18]; esto con el objetivo de determinar los diferentes algoritmos que existen

para el reconocimiento de imágenes aplicando técnicas de visión artificial, mediante la inclusión de artículos científicos originales realizados entre los años 2020 al 2022.

La recomendación PRISMA consiste en una lista de verificación y un diagrama de flujo de cuatro pasos, aplicados en la presentación de informes de revisiones sistemáticas y metaanálisis. Inicialmente, esta técnica se centró en los ensayos clínicos aleatorizados, pero PRISMA también se puede utilizar como base para informar revisiones sistemáticas de otros tipos de investigación y diversas áreas [19]. De igual forma, la recolección del contenido se ejecutó a través de la base de datos especializado IEEE Xplore; como motor de búsqueda a ScienceDirect; como bases de datos multidisciplinarias se eligieron Engineering Source, EBSCO y ProQuest.

Para conceptualizar mejor la búsqueda se utilizó palabras clave mediante operadores booleanos, obteniendo cadenas de consultas como: “Reconocimiento de imágenes” AND “Visión Artificial” OR “inteligencia artificial”; “Métodos” OR “Tecnología” OR “Componentes” AND “Reconocimiento de imágenes”, y cadenas de palabras clave en inglés como: “*Image Recognition*” AND “*Artificial Vision*” OR “*Artificial Intelligence*”; “*Methods*” OR “*Technology*” OR “*Hardware*” AND “*Image Recognition*”.

De los documentos encontrados, se seleccionaron aquellos artículos en lengua inglesa y castellana que respondan a la pregunta de investigación, de tal manera que pueda replicarse el estudio, en un rango de tiempo comprendido entre los años 2020 al 2022. Además, los artículos debían contener enfoques mayormente cualitativos o cuantitativos. Esos documentos se importaron al software de gestión de referencias Mendeley para su gestión en la eliminación de citas duplicadas, que luego de pasar por una revisión sobre aquellos que no cumplen con los criterios de selección luego de la revisión por tipo de estudio, resumen y texto completo.

Finalmente, para simplificar la evaluación y análisis de los datos extraídos se utilizó una matriz para organizar la información seleccionada de acuerdo con el nombre del autor, año de publicación, categorías de variable de estudio, tipo de estudio, tipo de documento científico y aportes.

III. RESULTADOS Y DISCUSIONES

En esta sección, se exponen los resultados de todos los pasos descritos de la revisión sistemática. Se han seleccionado los trabajos que compilan o proponen el desarrollo de técnicas o algoritmos para el procesamiento de imágenes mediante la visión artificial. Las segmentaciones del proceso de selección de datos se presentan en la Fig. 1. Además, se expone la distribución de los resultados hallados en la Tabla 1.

En la Fig. 2 se describen las tecnologías involucradas mediante el proceso de revisión. Las redes neuronales Artificiales (ANN) son una sucesión de capas que contienen un gran número de unidades interconectadas entre las llamadas neuronas [13]. Localidad espacial de los datos de entrada (SLID) es un método que permite la reducción de las operaciones de memoria y computación para lograr realizar un

modelo más eficiente [14]. De la misma forma, Machine Learning (ML) es una rama de la inteligencia artificial que puede ser usada en múltiples tareas con el fin de predecir alguna información futura [15].

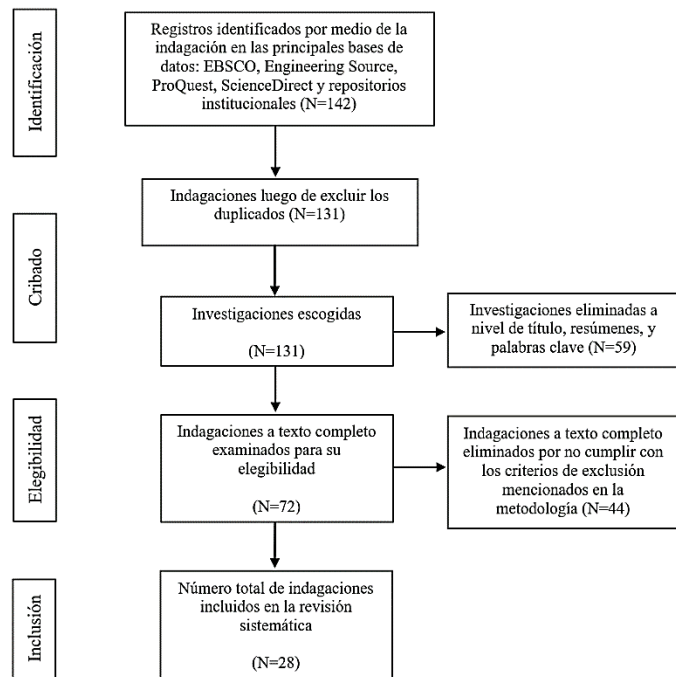


Fig. 1. Flujograma de selección de la declaración PRISMA. Fuente: creación propia.

Por otro lado, procesamiento de lenguaje Natural (NLP) es el método basado en inteligencia artificial para estudiar la manera de comunicar las máquinas con personas mediante lenguas naturales [16]. Convolution Neural Network (CNN) es una técnica conformada por una serie de capas convolucionales que se aplica para el reconocimiento de imágenes y aplicaciones por visión artificial [4]. De la misma forma, Deep Learning (DL) es una técnica que consiste en obtener una gran cantidad de datos y tras un número importante de entrenamiento lograr que el ordenador termine aprendiendo por cuenta propia [3]. Máquina de Vector Soporte (SVM) es una técnica que se emplea ampliamente en aplicaciones médicas para la clasificación de características completas en dos clases [17]. De igual forma, Red neuronal Binaria (BNN) es un método que nos permite refinar los pesos y activaciones al valor de un 1 bit aumentando el rendimiento del sistema y disminuyendo el consumo de recursos [11].

El método híbrido se trata de la implantación de dos técnicas diferentes con la finalidad de generar mejores resultados [18]. Por otro lado, Background Sustractor es una técnica muy utilizada en aplicaciones de detección de personas, donde se analiza cada uno de los objetos de la imagen y determina si corresponden al fondo o al cuadro de la imagen [19]. En el caso de Cascading detectors es una técnica que permite un filtrado progresivo para asignar un conjunto de características dado a una clase con la que se encuentra una

mayor similitud. Igualmente, redes Neuronales convolucionales cuánticas (QOCNN) es un método que implica tomar las circunvoluciones y la agrupación típica de las CNN y traducirlas de forma cuántica [20]. YOLO (You Only Look One) es una técnica de detección de objetos múltiples extremadamente rápida en tiempo real [21]. Finalmente, las redes Neuronales de Picos (SNN) es una técnica donde las neuronas se presentan en forma puntiaguda denominadas redes neuronales de tercera generación [22].

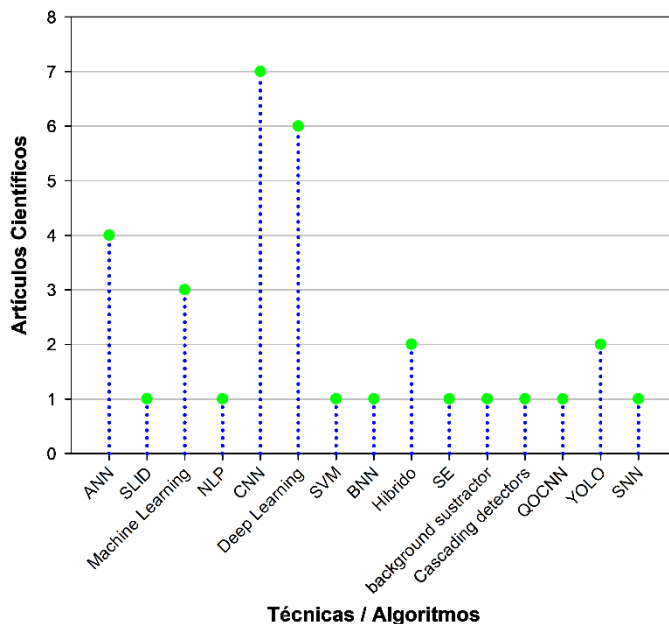


Fig. 2. Distribución de las investigaciones recuperadas por técnicas.

La Tabla 1 presenta la clasificación de los artículos seleccionados por autor(es), categorías, tipo de estudio, documento científico y aporte. El 46,43 % corresponde a la subcategoría tecnología (13 de las 28 investigaciones se correlacionan con las otras subcategorías). El 60,71 % de las investigaciones seleccionadas corresponden a la subcategoría aplicaciones (17 de los 28 muestran correlación con las otras subcategorías). El 32,14 % de las investigaciones seleccionadas corresponden a la subcategoría equipos (9 de los 28 muestran correlación con las otras subcategorías). El 100 % de las investigaciones corresponden a la subcategoría técnica (25 de las 28 investigaciones muestran correlación con las otras subcategorías). Finalmente, el 100 % de las investigaciones contenidas fueron de tipo cualitativo.

Al evaluar los resultados se encontró que la tecnología, equipos y técnicas fueron categorías centrales y determinantes en el desarrollo del reconocimiento de imágenes procesadas por visión artificial. Como se explica en [8], la creación de un nuevo conjunto de datos de entrenamiento, llamado FERPlus-A con base en una variedad de algoritmos de procesamiento de imágenes, se logró una mejora del rendimiento de generalización y una mayor precisión de la clasificación de la red neuronal convolucional.

TABLA I
COMPARATIVA DE LOS APORTES DE AUTORES REFERENTE A LOS TÓPICOS DE VISIÓN ARTIFICIAL

No.	Autor	CV				Tipo de estudio	Documento científico	Aportes
		Tecnología	Aplicaciones	Equipos	Técnicas			
1	[1]	X	X		X	Básico	Artículo de Investigación	Empleó las técnicas de Visión Artificial como redes neuronales artificiales y el algoritmo de codificación dispersa, aplicándolo al reconocimiento de manzanas dañadas por plagas y enfermedades mediante la tecnología del software MATLAB en condiciones controladas. Lograron un modelo del 90% de precisión.
2	[21]	X			X	Básico	Artículo de Investigación	Propone la técnica de datos de entrada de localidad espacial (SLID) para mejorar el rendimiento de los equipos con recursos limitados, soportado por la tecnología de capas alternas de LeNet y AlexNet. Consiguiendo una pérdida mínima de precisión y reduciendo la memoria RAM.
3	[30]		X		X	Básico	Artículo de Investigación	Examina diferentes estudios de inteligencia artificial que se enfocan en aplicaciones de discapacidad visual, optimizando el apoyo a personas que muestran baja visión. Además, se alcanza a averiguar que las técnicas de IA disponibles tratan sobre: Redes neuronales artificiales, aprendizaje automático, visión artificial y procesamiento del lenguaje natural.
4	[31]	X	X		X	Básico	Artículo de Investigación	Aprovechó la técnica de Visión Artificial como redes neuronales convolucionales para aplicarlo al reconocimiento de frutos de durazno maduros, así como frutos dañados, mediante la tecnología de las librerías de Keras y TensorFlow en condiciones controladas. Logrando un modelo del 95.31% de precisión a la hora de elegir entre frutos sanos y dañados.
5	[4]	X	X		X	Básico	Artículo de Investigación	Expone investigaciones enfocadas en las técnicas de inteligencia artificial específicamente aprendizaje automático (ML) y aprendizaje profundo (DL) aplicándolo a la detección de neoplasias tiroideas basándose en el análisis de imágenes médicas mediante la tecnología de la Radiómica para extraer características de las imágenes médicas para tener un conjunto de imágenes robusto.
6	[3]		X		X	Básico	Artículo de Investigación	Examina investigaciones en base a la técnica de aprendizaje profundo (DL) aplicándolo al reconocimiento automatizado de los síntomas de la retinopatía diabética presentes en la retina mediante la tecnología del conjunto de imágenes (Datasets) etiquetadas como ROC, E-Ophta, Kaggle DR, entre otros. De esta manera poder hacer bien el entrenamiento del modelo.
7	[7]		X	X	X	Básico	Artículo de Investigación	Estudia el uso de la técnica del algoritmo de la máquina de vectores de soporte (SVM) enfocado al reconocimiento de imágenes aplicado al procesamiento de paralelización de características del histograma de gradientes orientados (HOG), soportado en base a los equipos de unidad de procesamiento de gráficos (GPU), con ello asegura la precisión, acelera la velocidad de reconocimiento de imágenes digitales.
8	[11]		X	X	X	Básico	Artículo de Investigación	Presenta la arquitectura de la técnica de red neuronal binaria (BNN) con implicancias en un robot agrícola para la toma de decisión al aplicar agentes biológicos en los cultivos deseados. Además, presenta una integración con el equipo de una FPGA Xilinx Zynq UltraScall+ para aumentar la precisión.
9	[32]		X		X	Básico	Artículo de Investigación	Propone la aplicación del procesamiento de imágenes a la detección de enfermedades infantiles mediante la técnica del aprendizaje automático. Alcanzando con ello la identificación de los glóbulos blancos con una tasa de reconocimiento de hasta el 90 %.
10	[23]		X		X	Básico	Artículo de Investigación	Revisa los avances aplicados al reconocimiento y visualización del lenguaje de señas mediante las técnicas del aprendizaje profundo, entre ellos la arquitectura 3D-CNN enfocado a los problemas de reconocimiento de imagen/video y la red LSTM, utilizado debido a su eficacia con datos de series temporales.
11	[8]	X	X	X	X	Básico	Artículo de Investigación	Manejó la técnica de red neuronal convolucional cuantificada que solo usa aritmética entera, aplicado al reconocimiento de emociones faciales, mediante la tecnología del conjunto creado de datos de entrenamiento FERPlus-A, logrando la precisión del 86,58 % de clasificación. Siendo implementado en el equipo FPGA SoC Xilinx ZC706, alcanzando estrategias eficientes de paralelismo y almacenamiento en caché de datos.

12	[2]		X		X	Básico	Artículo de Investigación	Expone los avances recientes en el campo de la cirugía mediante las técnicas de aprendizaje profundo (DL) como las redes neuronales Convolucionales, aplicándolo a la ecografía, la tomografía computarizada, la resonancia magnética, entre otros.
13	[33]	X	X		X	Básico	Artículo de Investigación	Propuso un algoritmo propio fusionado con la técnica de redes neuronales convolucionales aplicado al reconocimiento y clasificación de imágenes de alta similitud, que en conjunto con la tecnología del conjunto de datos de imágenes de gemas y manzanas se consiguió que el algoritmo tenga una precisión mayor al 90%.
14	[24]		X		X	Básico	Artículo de Investigación	Planteó un sistema híbrido en base a la técnica de red neuronal convolucional (CNN) con los tipos de clasificadores tales como máquina de vector de soporte (SVM), k-vecino más cercano (KNN) y bosque aleatorio (RF); para la aplicación de detección automática de neumonía a partir de imágenes de rayos X de tórax, logrando una mejora en el tiempo de clasificación.
15	[12]				X	Básico	Artículo de Investigación	Los resultados obtenidos determinan que los modelos apoyados en técnicas de redes neuronales convolucionales realizan un mejor trabajo frente a otros modelos basados en Inteligencia Artificial. Las CNN han sido capaz de determinar una mejor detección y diagnóstico de lesiones en el territorio maxilofacial.
16	[34]		X		X	Básico	Artículo de Investigación	Los hallazgos demuestran un amplio panorama de técnicas o modelos que se aplican para la construcción de sistemas expertos para aplicaciones sobre la selección de granos de café mediante inteligencia artificial, mostrando ser efectivos y eficientes a más del 80% según los artículos analizados.
17	[5]	X		X	X	Básico	Artículo de Investigación	Gracias al desarrollo de la tecnología informática y la aparición de tecnologías emergentes del Deep Learning, hizo posible el reconocimiento de patrones mediante la técnica de redes neuronales artificiales, donde se utilizó una placa de desarrollo utilizada en este artículo es ARM9mini2440, que incluye una gran cantidad de dispositivos de hardware.
18	[10]	X		X	X	Básico	Artículo de Investigación	La técnica red neuronal convolucional ResNet en conjunto con la tecnología de la librería de Keras, lograron mejorar la velocidad de convergencia del entrenamiento y la precisión de la clasificación de imágenes de té verde, mediante el conjunto de equipos CPU Intel i7-7700hq.
19	[35]				X	Básico	Artículo de Investigación	Emplean técnicas de Deep Learning con redes neuronales de cuantiosas capas para la detección de rostros y reconocimiento de patrones. Incide que el porcentaje de exactitud de los algoritmos varían de acuerdo con la cantidad de datos del conjunto que han sido usados para ser entrenados.
20	[36]				X	Básico	Artículo de Investigación	Las técnicas híbridas y la red neuronal tipo Hopfield aportan mucho en el campo de la visión artificial, posibilitando el entrenamiento de nuevos patrones con novedosas formas. Particularmente en el caso de procesamiento de objetos en la identificación de poliedros tiene una eficiencia del 97.4%.
21	[9]		X		X	Básico	Artículo de Investigación	Concluyen que las técnicas del aprendizaje profundo han tenido un gran éxito en aplicaciones de reconocimiento de imágenes, reconocimiento de voz y traducción automática. Entre ellos, las máquinas de tensor de soporte han logrado avance en detección de imágenes, clasificación de imágenes, reconocimiento facial y otros dominios relacionados con la visión artificial.
22	[28]	X	X		X	Básico	Artículo de Investigación	Con la aparición de diferentes técnicas del Deep Learning se ha conseguido llevar a la visión artificial a otro nivel, facilitando el desarrollo de varias aplicaciones de visión por computador. En el sistema de detección de aves se obtuvo un 88% de precisión indicando que el modelo tiene un buen desempeño en la detección de aves utilizando tecnologías como Roboflow, YOLOv5 y Pytorch en Python, todas herramientas OpenSource.
23	[37]			X	X	Básico	Artículo de Investigación	Refiere que las soluciones basadas en equipos como el sensor de radio frecuencia, tienen la desventaja de implicar costos adicionales, mientras que algunas soluciones apoyadas en técnicas de redes convolucionales son más económicas y confiables. En este caso, se usó un conjunto de 65000 fotografías de huellas dactilares, obteniendo una exactitud de más del 99% demostrando la versatilidad de esta técnica.
24	[38]	X		X	X	Básico	Artículo de Investigación	Se implementó la tecnología YOLO junto con la transformada circular de Hough, logrando aumentar tanto la precisión como la exactitud a 67,9% y 80%, demostrando que si es posible mejorar el sistema de detección con base en estos parámetros gracias a las técnicas de inteligencia artificial entrenados para la detección de objetos. Sin embargo, el sistema no está diseñado para trabajar 24 horas, 7 días a la semana. Para mejorarlo, se requiere de equipo especializado.

25	[27]	X		X	X	Básico	Artículo de Investigación	Empleó una técnica red neuronal convolucional óptica cuántica (QOCNN), usando como tecnología el conjunto de datos MNIST en una computadora cuántica, aliviando el cuello de botella computacional en futuras aplicaciones de visión artificial.
26	[29]	X			X	Básico	Artículo de Investigación	Los resultados del experimento han demostrado que nuestros modelos SNN convertidos mediante la técnica de equilibrio de umbral alcanzan una precisión aún mayor que el modelo CNN entrenado correspondiente para un conjunto de datos de la tecnología MNIST con baja latencia. Además, para tecnológicas como Fashion-MNIST y CIFAR-10, nuestros SNN convertidos han mostrado menos pérdida de conversión que otros métodos en latencias bajas.
27	[26]		X	X	X	Básico	Artículo de Investigación	Las pruebas realizadas para las técnicas de sustracción de fondo y detectores en cascada para aplicaciones de detección de rostros en la plataforma de Windows mediante una computadora personal indica un tiempo de respuesta aproximado de 5 veces menor respecto al empleado en Linux por medio de la placa embebida Raspberry Pi 3B+. La técnica de sustracción de fondo requiere una memoria mayor al 26.28% disponible en el computador personal respecto a la técnica de detectores en cascada.
28	[6]	X			X	Básico	Artículo de Investigación	Se compara la tecnología de reconocimiento de IA con la de procesamiento de imágenes, centrándose en la aplicación de tecnologías de reconocimiento de inteligencia artificial en el procesamiento de imágenes. Se utiliza la técnica para realizar la detección, red neuronal Backpropagation, puesto que esta red muestra mejores resultados que el algoritmo tradicional en el experimento de simulación, sin embargo, tiene muchas deficiencias que deben mejorarse en el trabajo futuro.

Nota: Recopilación de investigaciones analizadas según los criterios de exclusión.

Del mismo modo, en [28] se enfatiza el uso de redes neuronales para la detección de patrones de rostros en imágenes, muestra una dificultad al realizar la clasificación entre imágenes que contiene caras y las que no. En tal sentido, se tiene como restricción de detección de rostros de manera frontal. Al evaluar los resultados, se encontró que la tecnología, equipos y técnicas fueron categorías centrales y determinantes en el desarrollo del reconocimiento de imágenes procesadas por visión artificial. Las mismas que, con el propósito de exponer una discusión lógica acorde con la finalidad de esta investigación, se argumentan a continuación.

Asimismo, empleando la técnica de red neuronal de retro propagación para la detección de patrones en imágenes de rostros, han logrado una eficacia del 97.1 % y un falso positivo del 2.8 %. Donde la arquitectura del sistema estaba compuesto por una red de 2 capas [12]. En contraste, en [17] sostiene que es mejor emplear el sistema híbrido de aprendizaje profundo para la detección y reconocimiento de objetos y rostros, puesto que alcanza un rendimiento muy eficiente con un tiempo de consumo de clasificación corto.

En cuanto a la categoría de tecnología, se debe precisar aquellos programas que hacen posible la integración de los algoritmos con la implementación y el tratamiento de los datos. Por ejemplo, en [21] se logró alcanzar un buen desempeño en el sistema de detección de aves relacionado con la precisión utilizando la tecnología *YOLO*. Esta tecnología permitió crear un algoritmo para el aprendizaje automático programado en Python. Mientras que la creación del conjunto de datos se realizó mediante *Roboflow*, una aplicación web fácil y amigable para el etiquetado y procesamiento de imágenes obtenidas.

Así mismo, en [26] se afirma que las tecnologías más usadas para tareas de visión artificial son *YOLO* y *Mask-RCNN* por sus ventajas de alta precisión y velocidad en comparación

con las tecnologías de reconocimiento tradicionales. De igual manera, en [28] se menciona que actualmente la mayoría de las tecnologías enfocadas a la visión artificial como *YOLO* hacen uso especialmente de la biblioteca *OpenCV*, librería que puede ser utilizada con lenguajes de programación como Python.

En cuanto a las aplicaciones, en [20], [2] y [8] afirman que la visión por computador se ha implementado en innumerables aplicaciones que van desde automóviles autónomos hasta los campos médicos de la radiología en el diagnóstico, aprovechando los marcos de aprendizaje profundo basados en redes neuronales convolucionales (CNN). La búsqueda de patrones abarca desde la predicción del riesgo cardiovascular a partir del análisis del fondo de ojo a la predicción del fallo renal agudo tras cirugía cardíaca.

Por otro lado, en [19] se argumenta que en aplicaciones de detección la técnica de sustracción de fondo toma relevancia. Ello se debe a que puesto que en las aplicaciones de detección de personas se requiere la separación del objeto del fondo de la imagen para posteriormente realizar procesos más complejos como rastreo y conteo. El alto número de cálculos que se efectúan en el entrenamiento y su elevada complejidad durante el entrenamiento de la red neuronal de los algoritmos de visión artificial hace que sea imprescindible un hardware o equipo de mayor capacidad de memoria RAM.

En [28], el porcentaje de exactitud de los algoritmos varían de acuerdo a la cantidad y calidad del conjunto de datos. En [20], se relata que el costo computacional de la IA está creciendo, puesto que *OpenAI* informa que la potencia computacional requerida para entrenar modelos modernos de IA se duplica cada 3.5 meses. Es por ello por lo que los ingenieros de IA han ido cambiando cada vez más a arquitecturas informáticas masivamente paralelas, como las

unidades de procesamiento gráfico o los chips aceleradores de IA personalizados.

IV. CONCLUSIONES

El objetivo de este trabajo fue identificar las técnicas, aplicaciones, equipos y tecnologías existentes aplicados en el reconocimiento de imágenes con visión artificial a través de una revisión sistemática de la literatura durante el periodo 2020 – 2022. Los resultados sugieren que las redes neuronales artificiales convolucionales, como las de tipo Hopfield, presentan una alta efectividad para realizar tareas de reconocimiento y clasificación de imágenes. Así, la combinación de herramientas tecnológicas como YOLO, Roboflow, Python y OpenCV muestran que el procesamiento de imágenes y el aprendizaje profundo están impulsando nuevas aplicaciones y herramientas que mejoran las diversas métricas de desempeño en estas tareas.

La visión artificial, a diferencia de las tecnologías que incorporan dispositivos electrónicos con sensores, permite una interpretación del entorno y ofrecen un mayor grado de representación de la realidad a partir de los procesamientos complejos de la información. La visión por computadora tiene diversas aplicaciones en el campo de la medicina, como para el diagnóstico de ecografía, tomografía computarizada y resonancia magnética, neoplasias tiroideas, retinopatía diabética, neumonía, entre otros.

Por lo tanto, la visión artificial, a diferencia de las tecnologías que incorporan dispositivos electrónicos con sensores, permite una interpretación del entorno con un mayor grado de representación de la realidad, siendo robustos a la complejidad del procesamiento de los datos. Sin embargo, la visión por computadora enfrenta retos en la necesidad de grandes demandas de memoria RAM en los equipos, lo que ha imposibilitado implementar una arquitectura más eficiente.

AGRADECIMIENTO

A la Universidad Privada del Norte (UPN) por el impulso promotor en este estudio de revisión sistemática. Al Dr. Eliseo Zarate-Perez por la orientación y gestión académica, muchas gracias.

REFERENCIAS

[1] Y. Abbaspour, A. Aghabara, M. Davari, and J. Maja, "Viabilidad del uso de técnicas de visión artificial e inteligencia artificial en la detección de algunas plagas y enfermedades de Apple," *Appl. Sci.*, vol. 12, no. 2, p. 906, Jan. 2022, doi: 10.3390/app12020906.

[2] D. Kitaguchi, N. Takeshita, H. Hasegawa, and M. Ito, "Visión artificial basada en inteligencia artificial en cirugía: avances recientes y perspectivas de futuro," *Ann. Gastroenterol. Surg.*, vol. 6, no. 1, pp. 29–36, Jan. 2022, doi: 10.1002/ags3.12513.

[3] J. Cuello, C. Barraza, and J. Escorcía, "Una revisión de los métodos de deep learning aplicados a la detección automatizada de la retinopatía diabética," *Rev. SEXTANTE*, vol. 23, no. 1, pp. 12–27, Dec. 2020, doi: 10.54606/Sextante2020.v23.02.

[4] F. Bini *et al.*, "Inteligencia artificial en el campo de la tiroides: una revisión exhaustiva," *Cancers (Basel)*, vol. 13, no. 19, p. 4740, Sep. 2021, doi: 10.3390/cancers13194740.

[5] G. Xu, "Diseño y Realización de Sistema de Reconocimiento Inteligente de Imágenes por Computadora," *J. Phys. Conf. Ser.*, vol. 2146, no. 1, p. 012002, Jan. 2022, doi: 10.1088/1742-6596/2146/1/012002.

[6] X. Zhang, "Aplicación de la Tecnología de Reconocimiento de Inteligencia Artificial en el Procesamiento Digital de Imágenes," *Wirel. Commun. Mob. Comput.*, vol. 2022, pp. 1–10, Jan. 2022, doi: 10.1155/2022/7442639.

[7] T. Fan, "Reconocimiento y Simulación de Imágenes Basado en Inteligencia Artificial Distribuida," *Complexity*, vol. 2021, pp. 1–11, Apr. 2021, doi: 10.1155/2021/5575883.

[8] J. Kim, J.-K. Kang, and Y. Kim, "Un acelerador CNN basado en FPGA solo para aritmética entera y eficiente en recursos para el reconocimiento de emociones faciales en tiempo real," *IEEE Access*, vol. 9, pp. 1–15, 2021, doi: 10.1109/ACCESS.2021.3099075.

[9] Z. Shi, Y. Ma, and M. Fu, "Fuzzy Support Tensor Product Clasificación adaptativa de imágenes para Internet de las cosas," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–11, Feb. 2022, doi: 10.1155/2022/3532605.

[10] H. Wang, K. Li, and C. Xu, "Una Nueva Generación de Modelo ResNet Basado en Inteligencia Artificial y Pocos Datos Impulsados y su Construcción en Modelo de Reconocimiento de Imágenes," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–10, Mar. 2022, doi: 10.1155/2022/5976155.

[11] C. Huang, "Un diseño de hardware/software basado en FPGA utilizando redes neuronales binarizadas para aplicaciones agrícolas: un estudio de caso," *IEEE Access*, vol. 9, pp. 26523–26531, 2021, doi: 10.1109/ACCESS.2021.3058110.

[12] P. Meléndez, C. Aráneda, and I. Tapia, "Revisión Sistemática Exploratoria: Diagnóstico asistido por computadora con uso de inteligencia artificial en imagenología maxilofacial," *Nac. Univ. Bello, Andrés Odontol. Fac. Mar, Viña*, vol. 5, pp. 1–21, 2020.

[13] D. Baldi, L. De Giorgis, M. Menini, F. Motta, and J. Colombo, "Efficacy of Instruments for Professional Oral Hygiene on Dental Implants: A Systematic Review," *Appl. Sci.* 2022, Vol. 12, Page 26, vol. 12, no. 1, p. 26, Dec. 2021, doi: 10.3390/AP12010026.

[14] N. Behnamnia, A. Kamsin, M. A. B. Ismail, and S. A. Hayati, "A review of using digital game-based learning for preschoolers," *J. Comput. Educ.*, pp. 1–34, Sep. 2022, doi: 10.1007/S40692-022-00240-0/TABLES/15.

[15] E. Zarate-Perez, E. Rosales-Asensio, A. González-Martínez, M. de Simón-Martín, and A. Colmenar-Santos, "Battery energy storage performance in microgrids: A scientific mapping perspective," *Energy Reports*, vol. 8, pp. 259–268, Nov. 2022, doi: 10.1016/J.EGYR.2022.06.116.

[16] E. J. Zarate Perez, A. Motta, and J. H. Grados, "Evolution of smart grid assessment methods: science mapping and performance analysis," *Curr. J. Appl. Sci. Technol.*, vol. 13, no. 12, pp. 5166–5175, Mar. 2021, Accessed: Jul. 23, 2021. [Online]. Available: <https://repositorio.upn.edu.pe/handle/11537/26745>

[17] E. Zarate-Perez, R. Sebastián, and J. Grados, "Online Labs: A Perspective Based on Bibliometric Analysis," in *19th LACCEI International Multi-Conference for Engineering, Education Caribbean Conference for Engineering and Technology: "Prospective and Trends in Technology and Skills for Sustainable Social Development" and "Leveraging Emerging Technologies to Con*, 2021, vol. 2021-July, doi: 10.18687/LACCEI2021.1.1.267.

[18] E. Zarate-Perez *et al.*, "Learning in Project-Based Engineering Education: A Bibliometric Analysis," *Proc. LACCEI Int. Multi-conference Eng. Educ. Technol.*, vol. 2022-July, 2022, doi: 10.18687/LACCEI2022.1.1.743.

[19] T. F. Galvão, T. de S. A. Pansani, and D. Harrad, "Principais itens para relatar Revisões sistemáticas e Meta-análises: A recomendação PRISMA," *Epidemiol. e Serviços Saúde*, vol. 24, no. 2, pp. 335–342, Jun. 2015, doi: 10.5123/S1679-49742015000200017.

[20] F. J. Valderrama-Purizaca, D. A. Chávez-Barturen, P. S. Muñoz-Pérez, V. ; Tuesta-Monteza, and H. I. Mejía-Cabrera, "Importancia de las redes neuronales artificiales en la ingeniería civil: Una revisión sistemática de la literatura Importance of artificial neural networks in civil engineering: A systematic review of the literature," *Rev. ITECKNE-Universidad*, vol. 18, no. 1, p. 2021, doi: 10.15332/iteckne.

[21] F. Alantali, Y. Halawani, B. Mohammad, and M. Al-Qutayri, "SLID:

- Explotación de la localidad espacial en los datos de entrada como método de reutilización computacional para CNN eficiente,” *IEEE Access*, vol. 9, pp. 1–9, 2021, doi: 10.1109/ACCESS.2021.3071409.
- [22] J. Anderson and H. Castro, “Aplicación de Machine Learning en la Gestión de Riesgo de Crédito Financiero: Una revisión sistemática,” *Interfases*, no. 015, pp. 160–178, Jul. 2022, doi: 10.26439/INTERFASES2022.N015.5898.
- [23] B. Joksimoski *et al.*, “Soluciones tecnológicas para el reconocimiento del lenguaje de señas: una revisión de alcance de las tendencias de investigación, desafíos y oportunidades,” *IEEE Access*, vol. 10, pp. 1–20, 2022, doi: 10.1109/ACCESS.2022.3161440.
- [24] I. Masad, A. Alqudah, A. Alqudah, and S. Almashaqbeh, “Un enfoque híbrido de aprendizaje profundo para construir un sistema inteligente para la detección de neumonía en la radiografía de imágenes de tórax,” *Int. J. Electr. Comput. Eng.*, vol. 11, no. 6, p. 5530, Dec. 2021, doi: 10.11591/ijece.v11i6.pp5530-5540.
- [25] W. F. Rafael Miño, P. V. R. Vilcherres Lizárraga, S. P. Muñoz Pérez, V. Tuesta Monteza, and H. I. Mejía Cabrera, “Modelamiento de procesos hidrológicos aplicando técnicas de inteligencia artificial: una revisión sistemática de la literatura,” *ITECKNE Innovación e Investig. en Ing. ISSN-e 2339-3483, ISSN 1692-1798, Vol. 19, N°. 1, 2022, págs. 46-60*, vol. 19, no. 1, pp. 46–60, 2022, doi: 10.15332/iteckne.v19i1.2645.
- [26] C. Niño, S. Castro, B. Medina, D. Guevara, J. Ramírez, and K. Puerto, “Comparación multiplataforma de técnicas basadas en visión artificial para detección de personas en espacios abiertos - Dialnet,” *Dialnet*, vol. 9, no. 1, pp. 22–33, Dec. 2021.
- [27] R. Parthasarathy and R. Bhowmik, “Red neuronal convolucional óptica cuántica: un nuevo marco de reconocimiento de imágenes para la computación cuántica,” *IEEE Access*, vol. 9, pp. 103337–103346, 2021, doi: 10.1109/ACCESS.2021.3098775.
- [28] L. Silva and J. Cobos, “Sistema de detección de aves mediante análisis de imágenes - Dialnet,” *Dialnet*, vol. 7, no. 6, pp. 1464–1484, Nov. 2021.
- [29] H. Viet and K. Myung, “Conversión efectiva de una red neuronal convolucional en una red neuronal de picos para tareas de reconocimiento de imágenes,” *Appl. Sci.*, vol. 12, no. 11, p. 25, Jun. 2022, doi: 10.3390/app12115749.
- [30] R. Alvarado and J. Llerena, “Revisión de la literatura sobre el uso de Inteligencia Artificial enfocada a la atención de la discapacidad visual,” *Rev. InGenio*, vol. 5, no. 1, p. 12, Jan. 2022, doi: 10.18779/ingenio.v5i1.472.
- [31] D. Arévalo, J. Ruiz, and J. Ayala, “Clasificación de frutos del durazno en maduros, no maduros y dañados hacia la cosecha automatizada,” *Dialnet*, vol. 10, no. 19, Jan. 2021.
- [32] J. Huang *et al.*, “Detección de enfermedades mediante tecnología de reconocimiento de imágenes de aprendizaje automático en inteligencia artificial,” *Comput. Intell. Neurosci.*, pp. 1–14, Apr. 2022, doi: 10.1155/2022/5658641.
- [33] Z. Liu, L. Sun, and Q. Zhang, “Algoritmo de clasificación y reconocimiento de imágenes de alta similitud basado en una red neuronal convolucional,” *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–10, Apr. 2022, doi: 10.1155/2022/2836486.
- [34] M. Monsalve, “Eficacia de sistemas expertos en la selección de granos de café (*Coffea arabica*),” *Rev. Amaz. Digit.*, vol. 1, no. 1, p. 13, Jan. 2022, doi: 10.55873/RAD.V1I1.163.
- [35] Q. Pinzón, S. Lazcano, H. Orozco, and M. Quintana, “Detección automatizada de emociones a través del rostro, una revisión del estado del arte,” *Coloq. Investig. Multidiscip.*, vol. 8, p. 1 a 8, 2020, doi: <http://hdl.handle.net/20.500.11799/112586>.
- [36] J. Ramírez, A. Rubiano, and R. Jiménez, “Object Recognition Through Artificial Intelligence Techniques,” *Rev. Fac. Ing.*, vol. 29, no. 54, p. e10734, Apr. 2020, doi: 10.19053/01211129.v29.n54.2020.10734.
- [37] J. Tellez, F. Lozano, L. Giraldo, and C. Higuera, “Detección de vida en huellas dactilares: una revisión,” *Elsevier*, p. 1 a 7, 2020, doi: <http://hdl.handle.net/1992/49299>.
- [38] J. Valencia, T. Ramirez, L. Castañeda, and T. Mauricio, “Detección de infracciones y matrículas en motocicletas, mediante visión artificial, aplicado a Sistemas Inteligentes de Transporte,” *SciELO*, no. 37, pp. 1–15, Jun. 2020, doi: 10.17013/RISTI.37.1-15.