

Sistema de Reconocimiento de Imágenes usando Servicios de Amazon Web Services

Image Recognition System using Amazon Web Services

Vanessa Mak-Huerta ^{a*}, Gabriel Antonio López-Valencia ^b, Indelfonso Rodríguez-Espinoza ^b, Santos Jesús Corral-Panduro ^c, Jazmín Alejandra Machado-Garrobo ^b

^a Departamento de Metal Mecánica, Tecnológico Nacional de México Instituto Tecnológico de Nogales, Avenida Instituto Tecnológico # 911, Col. Granjas. C.P. 84065, Nogales, Sonora, México.

^b Departamento de TICs, Universidad Tecnológica de Nogales, Avenida Universidad # 271, Col. Universitaria. C.P. 84097, Nogales, Sonora, México.

^c Departamento de Ciencias Económicas y Administrativas, Tecnológico Nacional de México Instituto Tecnológico de Nogales, Avenida Instituto Tecnológico # 911, Col. Granjas. C.P. 84065, Nogales, Sonora, México.

Correo electrónico: vanessa.mh@nogales.tecnm.mx

(Recibido: 20 de marzo 2025; Aceptado: 28 de abril 2025; Publicado: 01 de mayo 2025)

Resumen

El presente trabajo muestra detalladamente cual es el proceso para entrenar un sistema de reconocimiento de imágenes usando los servicios de Amazon Web Services. Los servicios en la nube que son proporcionados por AWS y Google son similares en cuanto a lo que ofrecen para el área de reconocimiento por visión artificial, en este caso se decidió probar los servicios de AWS para el reconocimiento de imágenes formados por conjuntos de entrenamiento con verduras y figuras geométricas básicas. La finalidad es conocer el proceso de entrenamiento de un sistema basado en Deep Learning como servicio de AWS, que resulte confiable y rápido en cuanto a los resultados esperados del reconocimiento de imágenes. Posteriormente, con el análisis de resultados poder desarrollar aplicaciones que sirvan a la industria de la manufactura y almacenes de exportación de frutas y verduras, así como también, a las instituciones educativas para que usen los avances tecnológicos en la enseñanza. Los resultados obtenidos muestran que es factible utilizar estas propuestas realizando ajustes experimentales para obtener mejores resultados en el proceso de reconocimiento.

Palabras claves: AWS, Artificial, Google, Reconocimiento, Sistemas, Visión.

Abstract

This work shows in detail the process to train an image recognition system using Amazon Web Services. The cloud services provided by AWS and Google are similar in terms of what they offer for the area of artificial vision recognition. In this case, it was decided to test AWS services for image recognition formed by training sets with vegetables and basic geometric figures. The purpose is to know the training process of a system based on Deep learning as an AWS service, which is reliable and fast in terms of the expected results of image recognition. Subsequently, with the analysis of results, we can develop applications that serve the manufacturing industry and export warehouses of fruits and vegetables, as well as educational institutions so that they can use technological advances in teaching. The results obtained shows that it is feasible to use these proposals by making experimental adjustments to obtain better results in the recognition process.

Keywords: AWS, Artificial, Intelligence, Recognition, Vision.

1. Introducción

El surgimiento de proveedores de servicios con técnicas de Inteligencia Artificial (AI) en la Nube ha creado la necesidad de empezar a conocer los procesos de uso para facilitar el entrenamiento de estructuras inteligentes basadas en Redes Neuronales Artificiales con algoritmos de entrenamiento de Machine y Deep Learning.

Amazon Web Services (AWS), ofrece diferentes servicios basados en algoritmos de inteligencia artificial para una diversidad de aplicaciones de reconocimiento específico, pero también ofrece procesos de entrenamiento con conjuntos de imágenes propias de acuerdo a las necesidades del usuario.

Estos procesos de entrenamiento pueden hacerse utilizando ambientes de programación en línea, como Google colab®, y en casos muy específicos de aplicación algunos de los diversos sistemas embebidos que actualmente existen. Uno de los objetivos de este trabajo de investigación es proporcionar las herramientas necesarias para que en un futuro puedan implementarse estos datos en sistemas embebidos con recursos pobres de procesamiento y dispositivos móviles. La informática en la nube es la distribución de recursos de TI bajo demanda a través de Internet mediante un esquema de pago por uso. En lugar de comprar, poseer y mantener servidores y centros de datos físicos, puede acceder a servicios tecnológicos, como capacidad informática, almacenamiento y bases de datos, en función de sus necesidades a través de un proveedor de la nube como es Amazon Web Services (AWS). Organizaciones de todo tipo, tamaño y sector utilizan la nube para una amplia gama de casos de uso, tales como respaldo de datos, recuperación de desastres, email, escritorios virtuales, desarrollo y prueba de *software*, análisis de Big Data y aplicaciones web con acceso para clientes. [1]

La electrónica se ocupa en prácticamente todos los ámbitos de la actividad humana como una forma de magnificar nuestros sentidos, como auxiliar para interpretar el mundo que nos rodea e incluso como punto de apoyo para, en cierta forma, incidir en él. Aunado a ello, se suma la red global que sirve para compartir información, sin importar el lugar y tiempo, y que, desde hace menos de una década, se entiende como el Internet de las Cosas (IoT) y que posibilita interconectar dispositivos electrónicos que hasta entonces estaban aislados. Al dotar a los objetos con esta capacidad, empieza a generarse información en enormes

cantidades que redimensionan las posibilidades de decisión y acción de la humanidad. Los sistemas embebidos, son parte central de este complejo mundo interconectado del IoT y con su Inteligencia Computacional conformada de *hardware* y *software*, cumplen con tareas determinadas en tiempo real que hacen que la inteligencia integrada en circuitos se multiplique día a día. [2].

Las Redes Neuronales son un subconjunto del aprendizaje automático, y están en el centro de los algoritmos de aprendizaje profundo. Están compuestas de capas de nodos, que contienen una capa de entrada, una o más capas ocultas y una capa de salida. Cada nodo se conecta a otro, y tiene un peso y umbral asociados. Si la salida de cualquier nodo individual está por encima del valor del umbral especificado, ese nodo se activa y envía datos a la siguiente capa de la red. De lo contrario, no se pasa ningún dato a la siguiente capa de la red. Existen varios tipos de redes neuronales que se utilizan para diferentes casos de uso y tipos de datos. Por ejemplo, las Redes Neuronales Recurrentes se usan comúnmente para el procesamiento del lenguaje natural y el reconocimiento de voz, mientras que las Redes Neuronales Convolucionales (ConvNets o CNN) se utilizan con mayor frecuencia para tareas de clasificación y visión artificial. [3]

El presente trabajo muestra una manera de trabajar con los servicios que proporciona AWS® en la nube, específicamente con el entrenamiento de Redes Neuronales con algoritmos de Deep Learning, de tal forma que los resultados puedan implementarse en trabajos subsecuentes en sistemas embebidos y se ajusten a los procedimientos de IoT.

2. Desarrollo

El uso de tecnologías de la Revolución Industrial 4.0., como la Visión por Computadora y la Inteligencia Artificial (IA) en el ámbito de la agricultura y la industria alimentaria. Especialmente, proporciona una comprensión integral de la visión y la inteligencia por computadora, metodologías que abordan diversas aplicaciones agrícolas como procesamiento de alimentos, aplicaciones basadas en agricultura, ganadería, datos de plantas análisis, riego inteligente y agricultura de próxima generación. [4]

En la construcción de arreglos de imágenes para entrenar redes con propósitos de clasificación, se debe tener alta variabilidad entre los elementos que lo

constituyen, de manera que, en su conjunto puedan dar una descripción lo más completa posible de la clase que se desea diferenciar. En el caso de los alimentos para diferenciar en este trabajo, se encontró que en aras de aumentar la eficiencia de los modelos a entrenar con las imágenes tomadas, es importante obtenerlas desde ángulos de manera que no se repitan las fotos cuando se haga aumento de datos, también tener en cuenta el entorno y las características de las imágenes que el modelo va a obtener una vez que está en funcionamiento, que para este trabajo son las formas más usuales en que las personas muestran los alimentos a la cámara, ayuda a aumentar la efectividad para la clasificación de objetos. [5]

Estudiar la posibilidad de incorporar el análisis espectral a la arquitectura CNN, de acuerdo con algunas arquitecturas reportadas en la literatura. Este novedoso sistema de clasificación, brinda un nuevo enfoque en la reestructuración de las capas convolucionales, la forma de generalizar el aprendizaje y la reducción del mapa de características. [6]

La implementación de una Red Neuronal Artificial para mejorar el reconocimiento de figuras geométricas que permita posteriormente, hacer estimaciones sobre formas en un contexto real para crear planos virtuales. Se utilizó el método de simulación, mediante el uso de un *software* informático, se simuló la actividad de una Red Neuronal para que pueda entrenarse y reconocer figuras geométricas. [7]

AWS dice que el uso de la informática en la Nube para el aprendizaje profundo hace posible incorporar y administrar fácilmente grandes conjuntos de datos que se utilizan para entrenar algoritmos, además de permitir a los modelos de aprendizaje profundo ajustar su escala de manera eficiente y reducir los costos con capacidad de procesamiento de GPU. Mediante el uso de redes distribuidas, la tecnología de aprendizaje profundo en la nube le permite diseñar, desarrollar y entrenar aplicaciones de aprendizaje profundo con mayor celeridad. [8]

3. Metodología

El presente trabajo muestra y propone el seguimiento de un proceso de AWS® para entrenar Redes Neuronales con Algoritmos de Machine Learning usando datos obtenidos mediante la aplicación de conceptos de Visión Artificial. Para esto, es necesario

que el desarrollador haga una cuenta en AWS® para que pueda utilizar algunos de los servicios que provee.

Es necesario también, que se instale en el dispositivo de cómputo del desarrollador un par de SDK's, para Python y boto3 de AWS®.

Para probar este proceso y poder demostrar la eficiencia y la conveniencia del uso de este tipo de herramientas de Inteligencia Artificial, se utilizaron 2 conjuntos de datos totalmente diferentes, ambos son datos definidos como imágenes contenidas en 3 diferentes clases, cabe indicar que el sistema se puede escalar a un número indefinido de clases, en la cual la cantidad depende del tipo de reconocimiento y clasificación que se requiera hacer.

Un conjunto de datos formado por 3 diferentes clases que contiene verduras, chile, tomate y cebolla. Un segundo conjunto definido por figuras geométricas, triángulo, círculo y rectángulo.

La razón de tomar estos datos para entrenar a la red es porque en la región existen almacenes que clasifican y empaacan diferentes tipos de verduras para el mercado de exportación y el estudio sobre las figuras geométricas es posible aplicarlo en el sistema educativo y poder reducir la curva de aprendizaje usando herramientas tecnológicas actuales.

Se utilizó el lenguaje de programación Python implementado en un Sistema Embebido Raspberry pi® para establecer la conexión a la Nube de AWS® y subir la imagen de la verdura o figura geométrica a reconocer.

Una vez que la imagen estaba en el *bucket* de almacenamiento de la nube se ejecutaban de manera automática las herramientas de AWS® para el reconocimiento de imágenes y generación del resultado. Los datos de entrenamiento y el conjunto de prueba se obtuvieron de manera manual mediante el uso de una webcam.

La figura 1 muestra el inicio del proceso de reconocimiento el cual permite entrar al ambiente de AWS® que valida ser un cliente de los servicios, permite crear un modelo de entrenamiento para Redes Neuronales Artificiales usando Algoritmos de Machine Learning y mediante la entrada de datos de imágenes poder definir la ejecución correcta del modelo de entrenamiento creado.

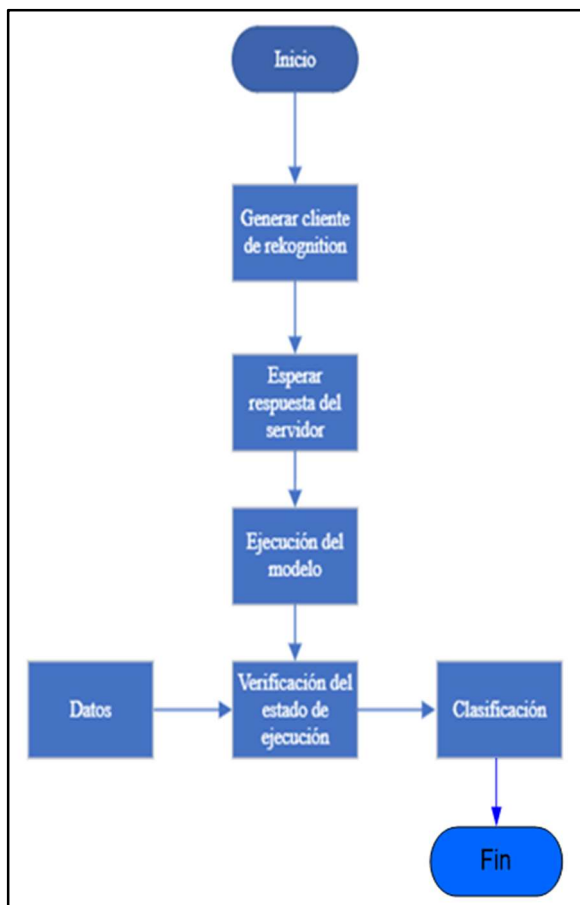


Fig. 1 Validación del modelo.

El proceso definido en el flujo mostrado en la figura 2, corresponde al análisis de los datos en tiempo real, este proceso puede ser alimentado con imágenes guardadas en alguna ruta del sistema de cómputo o pueden ser tomadas al instante mediante una cámara que en su versión más simple puede ser del tipo webcam.

Se le asigna a la imagen de prueba una etiqueta, que puede corresponder a una de las 3 clases propuestas en el estudio, del tipo verdura o figuras geométricas. Esta información entra al modelo entrenado previamente y se genera una etiqueta de reconocimiento que contiene la información de lo que es y el porcentaje con el cual el sistema afirma a que figura se parece más.

Finalmente, la figura 3 muestra el proceso que es imposible dejar de un lado ya que no detener el proceso de entrenamiento y reconocimiento en tiempo real genera gastos que pueden incrementarse a medida que el transcurra el tiempo de ejecución.

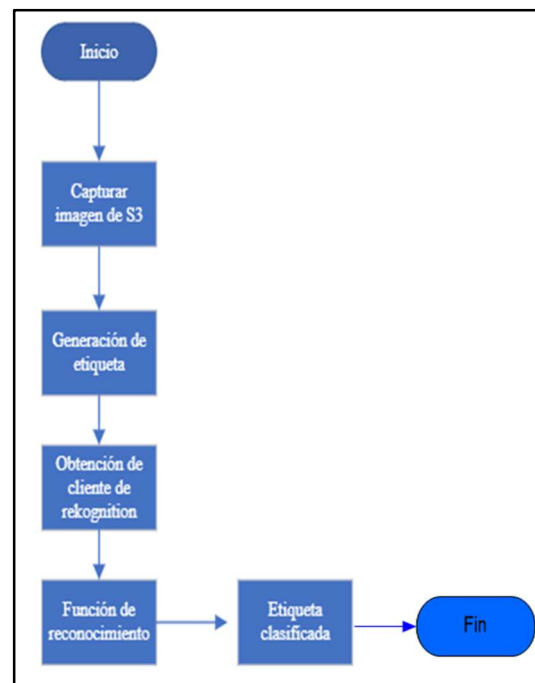


Fig. 2 Análisis de datos.

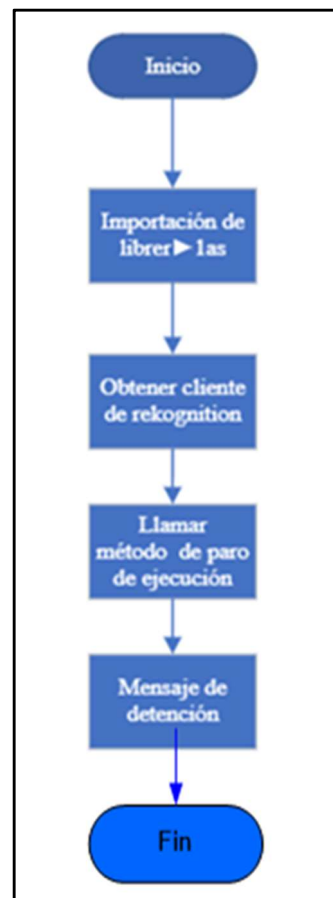


Fig. 3 Detención de ejecución del modelo.

4. Resultados y discusión

El sistema fue entrenado para reconocer figuras geométricas como son el triángulo, cuadrado y círculo, además, para reconocimiento de 3 verduras básicas, tomate, chile y cebolla.

Para validar los resultados del entrenamiento se utilizó el equivalente a un 10% del total de las imágenes usadas en el entrenamiento, esta cantidad de imágenes de prueba no formaron parte del conjunto de imágenes para entrenar a la red, es decir, se utilizó la técnica de validación cruzada. Los resultados se muestran en las tablas 1 y 2.

La tabla 1 nos muestra el porcentaje de reconocimiento de la red neuronal propuesta con respecto a las verduras, este análisis indica que las verduras tienen un alto porcentaje de reconocimiento positivo con aproximadamente un 1% de error.

Tabla 1 Porcentaje de reconocimiento de verduras.

Clase	Porcentaje de reconocimiento
Tomate	98.84 %
Chile	99.73 %
Cebolla	99.62 %

La tabla 2 muestra la precisión de la RNA propuesta, los resultados indican que cada vez que se validaba el reconocimiento de una de las verduras, la reconocía tal como era, es decir, cada vez que se le presentaba un tomate el sistema lo reconocía como tal y así sucesivamente con el chile y la cebolla, no hubo confusión por parte del sistema de RNA para clasificar a cada una de las verduras.

Tabla 2 Porcentaje de reconocimiento de verduras.

Clase	Precisión en el reconocimiento
Tomate	100
Chile	100
Cebolla	100

La tabla 3 únicamente muestra el formato de matriz de confusión utilizada para definir el grado de eficiencia de la RNA propuesta, se utilizaron las siguientes abreviaciones:

- **VP:** Es una verdura correcta dentro del conjunto de pruebas y el sistema la reconoce de manera adecuada.
- **VN:** El sistema confunde, por ejemplo, una verdura con otra
- **FP:** Se presenta al sistema una verdura real y no la identifica correctamente
- **FN:** No es una verdura real y el sistema no la reconoce como tal.

Los datos numéricos se muestran en las tablas 4 y 5.

Tabla 3 Matriz de confusión para reconocimiento propuesta.

	Predicción	
Real	0	1
0	FN	VN
1	FP	VP

La tabla 4 muestra como de los 45 datos usados como vectores de prueba, todos fueron identificados de manera correcta, esto lo podemos observar en la intersección de los resultados de la predicción con lo que realmente es.

Tabla 4 Matriz de confusión para reconocimiento de verduras.

	Predicción	
Real	0	1
0	0	0
1	0	45

El 1 indicado en el parámetro *real* muestra que es una verdura válida dentro del conjunto de pruebas y el 1 en el parámetro de la *predicción* indica que se reconoció de forma correcta, VP. Es decir, de las 45 verduras mostradas al sistema, las reconoció correctamente.

No hubo ningún dato identificado de forma incorrecta, es decir un falso positivo, VN, esto es, por ejemplo, si se presentó una imagen de chile, esta no fue reconocida como cebolla o tomate.

De igual manera, no hay datos catalogados como falsos negativos, FN y verdaderos negativos, VN. La exhaustividad se calcula mediante:

$$Exhaustividad = \frac{45}{45 + 0} = 1$$

Este resultado mostrado como exhaustividad indica que la RNA propuesta está clasificando de manera acertada cualquiera de las verduras propuestas.

Con respecto a las figuras geométricas, la tabla 5 muestra como el porcentaje de reconocimiento es menor con respecto a las verduras, muestra una inestabilidad en el proceso de reconocimiento ya que el rectángulo es de un 95 %, las otras 2 figuras andan entre 65 y 70 %.

Lo que representa que la RNA propuesta no es una buena opción para esta tarea.

Tabla 5 Porcentaje de reconocimiento de figuras geométricas.

Clase	Porcentaje de reconocimiento
Triángulo	70.93 %
Círculo	64.97 %
Rectángulo	95.42 %

La tabla 6 muestra también, como la RNA tiene problemas para reconocer de forma adecuada y consistente a cada una de las figuras geométricas propuestas, además confirma la inestabilidad indicada en la tabla 5, ya que es muy precisa para reconocer el rectángulo y el resto de las figuras no.

Tabla 6 Precisión del reconocimiento de figuras geométricas.

Clase	Precisión en el reconocimiento
Triángulo	71.42
Círculo	57.14
Rectángulo	100

La tabla 7 muestra el formato de matriz de confusión para reconocimiento de figuras propuesta, utilizada para

definir el grado de eficiencia de la RNA propuesta, se utilizaron las siguientes abreviaciones:

- **VP:** Es una figura geométrica correcta dentro del conjunto de pruebas y el sistema la reconoce de manera adecuada.
- **VN:** El sistema se confunde, se le presenta una imagen y la clasifica correctamente
- **FP:** Se presenta al sistema una figura real y no la identifica correctamente
- **FN:** No es una figura real y el sistema no la reconoce como tal.

Tabla 7 Matriz de confusión para reconocimiento de figuras propuesta.

	Predicción	
Real	0	1
0	FN	VN
1	FP	VP

La tabla 8 muestra como de los 21 datos usados como vectores de prueba, 16 fueron identificados de manera correcta, hubo 5 datos identificado de forma incorrecta, es decir 5 falsos positivos. De igual manera, no hay datos catalogados como falsos negativos y verdaderos negativos. El resultado del parámetro de exhaustividad nos indica que esta RNA no es una propuesta adecuada para la clasificación de figuras geométricas debido al bajo valor porcentual que muestra.

Tabla 8 Matriz de confusión para reconocimiento de figuras geométricas.

	Predicción	
Real	0	1
0	0	0
1	5	16

La exhaustividad es:

$$Exhaustividad = \frac{16}{16 + 5} = 0.76$$

5. Conclusiones

El uso de las herramientas de Inteligencia Artificial que muestra AWS en cuanto al reconocimiento usando datos obtenidos mediante Visión Artificial es adecuado para identificar

características de diversos objetos, como lo muestran los resultados obtenidos en el presente trabajo. Una de las ventajas principales que se muestran en este desarrollo es el poder trabajar con los procesos de entrenamiento de una Red Neuronal en la Nube, donde se pueden minimizar la gran cantidad de cálculo que se ocupa para el entrenamiento con algoritmos complejos de Machine Learning.

Las tablas 1 y 2 muestran el comportamiento de la red neuronal entrenada para clasificar verduras, podemos observar que el porcentaje de reconocimiento es muy alto, presentando un porcentaje de error menor al 2%, lo que lo hace una propuesta interesante para su aplicación en líneas de producción de los almacenes de procesamiento de verduras.

Las tablas 3 y 4 muestran resultados del proceso de reconocimiento de figuras geométricas en las cuales se puede observar que en 2 diferentes clases, círculos y triángulos, el porcentaje de reconocimiento y precisión es bajo, no así en las figuras correspondientes a los rectángulos. El análisis de estos resultados puede explicar el comportamiento numérico y se debe principalmente a 2 características principales en una imagen: resolución y proceso de dibujo. El tamaño de las imágenes es de 200 x 200 pixeles, lo cual es muy baja si las comparamos con los datos usados para la clasificación de las verduras, además, el proceso de dibujo es totalmente manual.

En general, el presente trabajo muestra una propuesta muy interesante de cómputo inteligente en la nube para aplicaciones de reconocimiento con técnicas de Inteligencia Artificial, lo cual repercute en que ya no dependamos de equipos con capacidades de procesamiento inferior fijos, como computadoras personales o sistemas embebidos.

Trabajos futuros implican el uso de sistemas embebidos como plataformas de tránsito de datos hacia la nube y como dispositivos de control de ejecución de estos algoritmos de reconocimiento en tiempo real.

6. Agradecimientos

Agradecemos al Tecnológico Nacional de México campus Nogales por todas las facilidades prestadas para el desarrollo de este trabajo.

7. Referencias

- [1] Amazon, «aws.amazon.com,» 2025. [En línea]. Available: https://aws.amazon.com/es/what-is-cloud-computing/?nc1=f_cc. [Último acceso: 2025].
- [2] F. Rodríguez, «infotec.mx,» 2025. [En línea]. Available: <https://www.infotec.mx/sistemas-embebidos-smart-la-era-del-internet-de-las-cosas>. [Último acceso: 2025].
- [3] IBM, «ibm.com,» 2025. [En línea]. Available: <https://www.ibm.com/mx-es/think/topics/convolutional-neural-networks>. [Último acceso: 2025].
- [4] Vijay Kakani, Van Huan Nguyen, Basivi Praveen Kumar, Hakil Kim, Visweswara Rao Pasupuleti. “A critical review on computer vision and artificial intelligence in food industry”. Department of Biomedical Sciences & Therapeutics, Faculty of Medicine and Health Sciences, University Malaysia Sabah. Elsevier B.V. 2020.
- [5] Muñoz Bocanegra Ricardo. “Aprendizaje profundo en dispositivo portable para el reconocimiento de frutas y verduras”. Universidad Autónoma de Occidente. Departamento de automática y electrónica. Santiago de Cali. 2019.
- [6] Fortuna-Cervantes Juan Manuel, Ramírez-Torres Marco Tulio, Mejía Carlos Marcela, Murguía Ibarra José Salomé, Martínez-Carranza Juan. “Sistema para clasificación de texturas en imágenes mediante aprendizaje profundo y características wavelet”. Universidad Autónoma de San Luis Potosí, Facultad de Ciencias-IICO. Research in Computing Science 150(6), 2021.
- [7] Van Adrianzén Olano, Juan Elias Villegas Cubas, Hugo Vega Huertas, Gisella Luisa Elena Maquen Niño. “Aplicación de una Red Neuronal Artificial en el reconocimiento de figuras geométricas”. Universidad Nacional Toribio Rodríguez de Mendoza. Universidad, Ciencia y Tecnología, vol.26, Núm. 114, (pp. 93-107). 2022.
- [8] Amazon, «aws.amazon.com,» 2022. [En línea]. Available: <https://aws.amazon.com/es/what-is/deep-learning/>. [Último acceso: 2025].
- [9] Lins, S., Konstantin D., Pandl., Teigeler, H., Thiebes, S., Bayer, C., Sunyaev, A. Artificial Intelligence as a Service. Received: 8 April 2020 / Accepted: 25 May 2021 / Published online. (2021).

[10] Cárdenas-Sánchez, B. C., & Olarte-Rojas, C. A. Análisis de seguridad entre microservicios con Amazon Web Service. *Revista Logos Ciencia & Tecnología*, 14(2), 42-52. (2020).

[11] Khrijil, S., Benbelgacem, Y., Chéour, R., Houssainil, D., Kanoun, O. Design and implementation of a cloud-based event-driven architecture for real-time data processing in wireless sensor networks. Accepted: 17 June 2021 / Published online: 26 July 2021 © The Author(s). (2021).

[12] Teja Swaroop. "Create your own face recognition Service with AWS rekognition". <https://youtube/oHSesteFK5c>. 2022.

[13] Khrijil S Bembengacem, Chéour R, Houssainil D, Kanoun O. "Design and implementation of a cloud based event driven architecture for real time data processing in wireless sensors networks". 26 Julu 2021.