

Mejorando la red neuronal convolucional para detectar objetos en imágenes infrarrojas

Daniel **Treviño Sánchez**
Vicente **Alarcón Aquino**

Desde que se propusieron por primera vez las redes neuronales convolucionales (*Convolutional Neural Network*, CNN, por sus siglas en inglés) se ha creado un gran número de aplicaciones para el análisis y procesamiento de las imágenes, como la detección de objetos en tiempo real. También se han realizado muchas investigaciones para aumentar la precisión y reducir el tiempo de ejecución de las CNN. Sin embargo, para mejorar el desempeño de las CNN es primordial poder extraer las características clave. Es por eso que identificar y extraer rasgos distintivos sigue siendo un campo abierto para la investigación. Inicialmente, para poder identificar los elementos más relevantes, es necesario conocer y analizar el tipo de imagen que se va a trabajar como, por ejemplo, las imágenes infrarrojas (IR).

RAYOS INFRARROJOS

La visión del ojo humano está limitada a lo que llamamos luz visible, una pequeña fracción del espectro de ondas que ronda el 1 % y se encuentra ubicada entre los rayos ultravioleta (UV) y los infrarrojos (IR). Ninguno de los dos es visible por el ojo humano, pero los segundos pueden ser percibidos como calor, aunque abarcan un rango mucho

más amplio, como se aprecia en la Figura 1. Su nombre proviene de la palabra latina *infra*, que significa debajo, ya que la banda IR está ubicada inmediatamente debajo de la banda de la luz roja.

Todos los objetos emiten rayos IR a temperaturas superiores al cero absoluto, es decir a $-273\text{ }^{\circ}\text{C}$, y la cantidad de radiación aumenta con la temperatura

(Larkin, 2018). La radiación IR se puede percibir utilizando cámaras especiales que convierten las variaciones térmicas a lo largo de los objetos, tomando el fondo de la escena como referencia, en una imagen visual. Este proceso se llama termografía y permite realizar mediciones de la temperatura de un objeto sin ningún contacto físico. El mayor uso para estas cámaras es la vigilancia; por ejemplo, la cámara diurna y nocturna puede usar tanto luz visible como con rayos IR. Este dispositivo produce un video a color durante el día y uno en escala de grises durante la noche, incluso en completa oscuridad. Sin embargo, la capacidad de ver tanto con la luz del día como en completa oscuridad no es la única ventaja de la tecnología IR. Adicionalmente, puede encontrar objetos escondidos en la niebla o debajo de un camino que podrían ser un riesgo, incluso puede ver a través de ciertos materiales, como el plástico o la tela, que la luz visible no puede.

DIFERENCIAS ENTRE LA VISIÓN IR Y LA VISIÓN NOCTURNA

Es importante diferenciar entre la visión IR y la visión nocturna, ya que no funcionan de la misma manera y por lo tanto no presentan las mismas ventajas. En primer lugar, la visión nocturna recolecta la poca cantidad de luz disponible y la amplifica para que pueda ser percibida por el ojo humano. Esto es

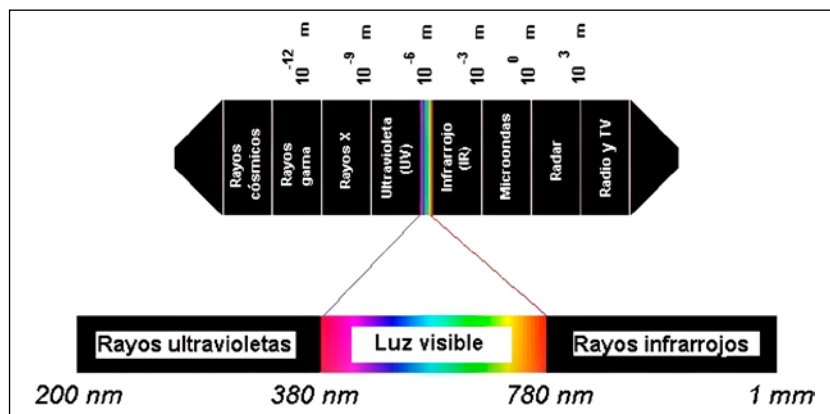


Figura 1. La luz visible dentro del espectro electromagnético. Todas las medidas están en nanómetros (metros $\times 10^{-9}$) (Larkin, 2018).

posible debido al hecho de que la luz está formada de pequeños paquetes de energía llamados fotones, que están presentes aun durante la noche, aunque en menor medida. Sin embargo, puede haber situaciones de total oscuridad en que falten dichos paquetes lumínicos o sean muy pocos; en estos casos, la visión nocturna resulta inútil. Contrariamente, la tecnología IR se basa en ondas térmicas emitidas por los objetos cuando disipan el calor. La diferencia de temperatura entre los diversos objetos proporciona el contraste para identificarlos. En otras palabras, los objetos emiten su propia luz IR en cantidades que varían según la cantidad de calor que tenga el cuerpo. Por lo tanto, los lentes IR se pueden usar en total oscuridad. Otra ventaja de la visión IR sobre la nocturna es que la primera puede percibir objetos parcial o totalmente ocultos (camuflajeados), ya que no depende de sus similitudes en forma o textura con el medio ambiente, sino de las diferencias térmicas que los contrastan. Por ejemplo, una persona escondida dentro de una caja de cartón o detrás de unos arbustos puede ser difícil de ver aun con visión nocturna. Sin embargo, la tecnología IR ve claramente su firma de calor que atraviesa el material de cobertura. La profundidad con la que el IR puede ver el material depende de la sensibilidad de la unidad.

Finalmente, la visión IR puede detectar objetivos cálidos a grandes distancias. Una visión térmica de última generación puede detectar animales de sangre caliente en un campo a más de 1,000 metros

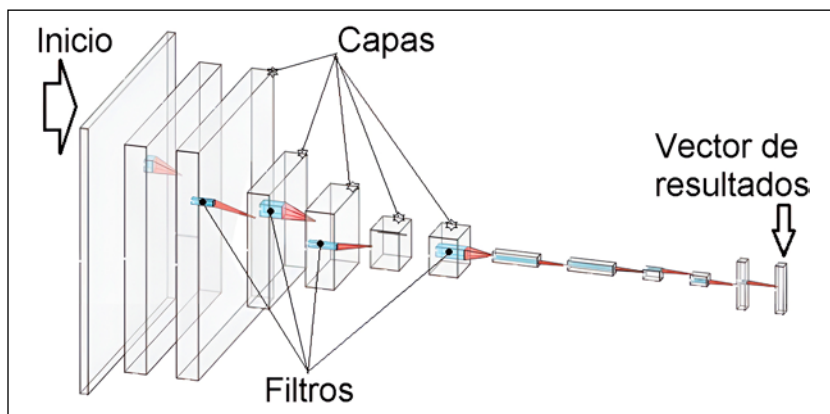


Figura 2. Ejemplo de una red neuronal convolucional.

de distancia (Coffey, 2011), aunque a esa distancia no es posible saber si se trata de un ciervo, una vaca o un cerdo. Sin embargo, al percibir la presencia y ubicación de dichos objetos, basta con acercarse un poco para confirmar el objetivo. Existen diversos tipos de tecnología óptico-térmica, pero entre mayor sea su calidad y sensibilidad, mayor será también su costo.

REDES NEURONALES CONVOLUCIONALES

Las CNN es un método muy usado para la detección de patrones que existe hace más de 50 años. Sin embargo, en años recientes se han hecho desarrollos importantes en el área y se han convertido en una de las herramientas más importantes en reconocimiento de imágenes, clasificación y localización de objetos.

Para diseñar una CNN es importante considerar el número de parámetros y operaciones matemáticas requeridas en cada capa, así como saber su operación en detalle. El proceso de clasificación de la CNN emplea una imagen como entrada, luego la etiqueta bajo una categoría predefinida (automóvil, persona, perro, tigre, etc.), como se muestra en la Figura 2. La imagen de entrada es una matriz de valores, cada uno representa un pixel. Dependiendo de las dimensiones de la imagen, serían alto (h) x ancho (w) x canales (c). Por ejemplo, una imagen con un arreglo $7 \times 7 \times 3$ sería una imagen a color en formato RGB (combinación de colores rojo,

verde y azul) y una imagen con un arreglo $7 \times 7 \times 1$ sería una imagen en escala de grises.

Antes de poder usar el potencial de una red neuronal, esta debe de ser entrenada para que aprenda a hacer lo que se desea de ella. Para ello es necesario contar con una base de datos (o imágenes) previamente etiquetados que le permitirán optimizar sus

parámetros, hasta lograr que su salida iguale el valor de las etiquetas. Dentro de la red convolucional existen diferentes tipos de capas; la más conocida es la convolución, que es precisamente la que le da nombre a la red. La función de esta capa es extraer características de la información de entrada. Para ello utiliza diferentes filtros según las características que se desee extraer. Estos filtros son una serie de matrices cuadradas de diferentes tamaños que actualizan sus valores durante la etapa de entrenamiento. Un solo filtro multiplica cada elemento de una región de la imagen y los suma para formar un elemento de salida; para obtener todos los elementos de salida el filtro se desliza por toda la imagen, como se muestra en la Figura 3. El resultado de tales operaciones se conoce como “mapa de características”. En las primeras capas de la red, la diversidad de filtros permite operaciones que detectan bordes, colores o esquinas, mientras que en las capas intermedias se encuentran patrones o se detectan características específicas. El segundo tipo de capa más usada es la de pooling, y se utiliza para reducir las dimensiones de los datos de entrada. Este método resume una región completa a un solo valor mediante una selección o un cálculo estadístico. Los métodos de pooling más conocidos son el pooling máximo y el pooling promedio. El entrenamiento de la red se realiza una vez que esta se ha armado y conectado correctamente, y la base de

datos se ha revisado y etiquetado adecuadamente. Como se mencionó anteriormente, el propósito de este proceso es encontrar los valores óptimos de todos los parámetros o filtros dentro de la red para producir las salidas deseadas al alimentarlo con datos similares a los del entrenamiento, pero que analiza por primera vez. Todos los

patrones clave se aprenden durante esta etapa, y cada parámetro se modifica muchas veces hasta encontrar su valor óptimo. Cuantos más patrones relevantes identifique y extraiga la red, mayor será su precisión para detectar e identificar objetos.

ANÁLISIS MULTIRRESOLUCIÓN

Uno de los principales beneficios del análisis multirresolución (*Multi-resolution Analysis*, MRA, por sus siglas en inglés) es que no pierde ninguna información al momento de descomponer los datos. Estos quedan contenidos en cuatro bloques llamados coeficientes. En el caso de las imágenes, el bloque principal contiene la mayor parte de la información, incluyendo la estructura y algunos patrones. Los otros tres contienen los detalles en horizontal, vertical y diagonal. Dependiendo de la ondita (wavelet) que se utilice, se pueden resaltar características específicas de la imagen, facilitando el trabajo de la red convolucional. En este sentido, existe un análisis multirresolución de segunda generación llamado esquema de elevación o *lifting scheme* (Sweldens, 1998), que se realiza mediante bloques de procesamiento. Estos bloques pueden integrarse como capas de operaciones matemáticas dentro de la red convolucional. Esta técnica de MRA preserva la información de frecuencia y espacio al reducir la dimensionalidad sin ninguna pérdida, incluso si solo se utilizan los coeficientes de aproximación.

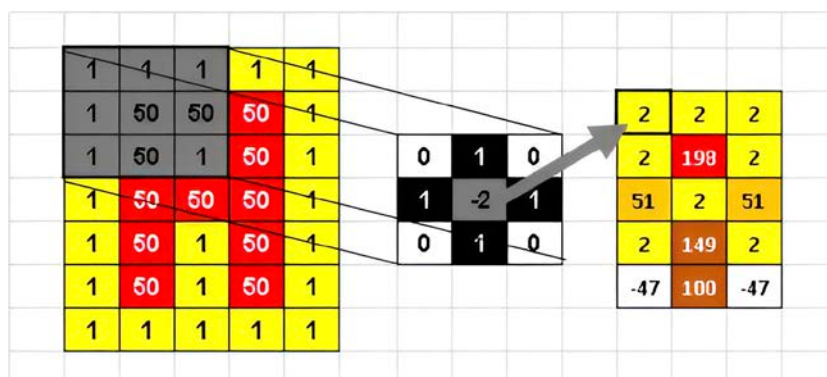


Figura 3. Ejemplo de una capa convolucional con su núcleo o kernel.

Estos coeficientes contienen la mayor parte de la información de energía y estructura de la imagen, que es la más parecida a la imagen original. Al hacer esto se pueden extraer características relevantes de las imágenes que permiten mejorar el desempeño de la CNN.

LIMITACIONES DE LAS IMÁGENES INFRARROJAS

A pesar de las bondades de los rayos IR sobre la luz visible, e incluso sobre la visión nocturna, las imágenes IR presentan algunos retos para poder manejar su contenido. En general, este tipo de imagen es muy susceptible al ruido térmico. Además, estas imágenes requieren que se acentúen los bordes y se aumente el contraste para que los objetos puedan reconocerse con mayor facilidad; por todo lo anterior se requieren varios filtros para ajustar la imagen. Por otro lado, a diferencia de una imagen de luz visible, que presenta tres canales para formar el color (rojo, verde y azul), las imágenes IR solo cuentan con un canal de escala de grises. Finalmente, el reto más grande es que las imágenes IR no cuentan con texturas o patrones que se puedan usar para reconocer los objetos. Las superficies de cada objeto radian según su temperatura, pero la imagen IR se forma con las diferencias térmicas que existen entre los objetos y el fondo. Una misma superficie puede tener diferentes valores térmicos en distintos puntos; por ejemplo, el patrón de temperaturas de un auto es muy distinto cuando está en reposo con el motor apagado que

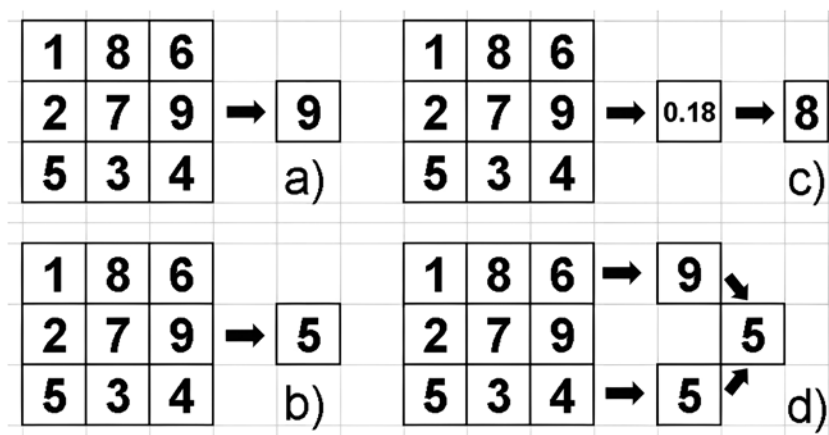


Figura 4. Ejemplos de cuatro tipos de poleo, a) poleo máximo, b) poleo promedio, c) poleo estocástico y d) poleo mixto.

cuando está en operación. Esto último no ayuda en la obtención de características de un objeto, pero sí da otro tipo de información sobre su estado. Para resolver este reto se requieren múltiples pasos orientados a acondicionar las imágenes IR, entre ellos el filtrado, el aumento de contraste y el resaltado de bordes.

POLEOS CONVENCIONALES

Las capas de poleo convencional más usadas son la de poleo máximo y poleo promedio (Lee *et al.*, 2018); en ambas capas se reduce toda una región a un solo valor. En el caso del poleo máximo, como su nombre lo indica, toma el valor más grande de todos los que forman la región y los demás valores se eliminan. Aunque con el poleo máximo se pierde información, la salida aún logra mantener algunos detalles de frecuencia, dependiendo del comportamiento de la señal de entrada. Este fenómeno permite que la CNN mejore la precisión, razón por la cual el poleo máximo fue seleccionado para este estudio sobre otros métodos convencionales. En el caso del poleo promedio, se pierde información relevante debido a que actúa como un filtro pasa-bajas, generalizando las características al obtener su valor medio de toda la región. Como se ve en la Figura 4a y 4b, respectivamente. Aunque existen otras técnicas de poleo más complejas, en general, después de seleccionar un valor para cada

región, los demás se eliminan. Por ejemplo, el poleo estocástico (Zeiler y Fergus, 2013) que toma el valor probabilístico de cada elemento de la región para seleccionar el valor de salida. Las probabilidades se calculan con base en la magnitud de cada elemento, siendo la unidad el valor sumado de todas las probabilidades de una región. Así, los valores más grandes tienen mayor probabi-

lidad de ser seleccionados que los valores pequeños; sin embargo, estos aún pueden ser escogidos de vez en cuando.

Otro método aleatorio menos complicado es el poleo mixto (Yu *et al.*, 2014); este simplemente toma el valor máximo y el valor promedio para seleccionar uno solo. El proceso funciona asignando una probabilidad del 50 % a los dos valores, y como si se lanzara una moneda al aire, se selecciona al ganador, como se puede ver en la Figura 4c y 4d, respectivamente. Los diferentes tipos de poleo seleccionan un valor que represente toda una región y simplemente borran toda la demás información. Esto no garantiza en ningún caso que se seleccione el mejor valor o el valor más adecuado para que la red aprenda las características esenciales o los patrones más útiles.

Por otro lado, la razón de que estos métodos sean tan usados es porque, al ser simples, facilitan el procesamiento en la red. Además, algunos de estos métodos son muy efectivos para prevenir el sobre-entrenamiento (*overfitting*) debido a su comportamiento no lineal, pero el resultado final depende totalmente del desempeño del resto de las capas. Por estas razones, son métodos que se pueden mejorar, pero a la vez tienen elementos que se pueden seguir aprovechando en los nuevos modelos.

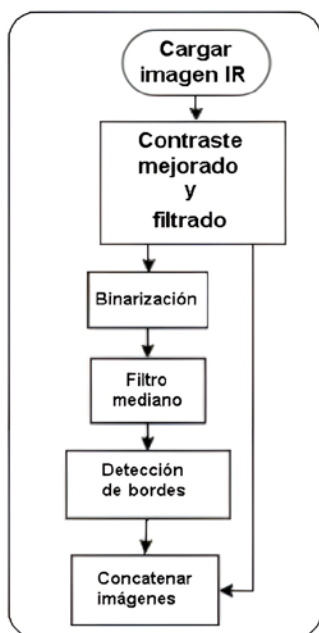


Figura 5. Preprocesamiento propuesto para mejorar la calidad de las imágenes IR y facilitar su uso por la CNN.

ANTECEDENTES

Existen algunos modelos que combinan el MRA con la CNN, como el trabajo de Bastidas-Rodríguez y colaboradores (2020), que utilizan un esquema de elevación adaptable que aprende la configuración de las onditas incorporando una pequeña CNN como parte de sus bloques internos. Luego, este modelo se integra en una CNN más grande para detectar escenas y clasificar texturas. Otros enfoques (como Williams y Li, 2016) también detectan escenas y clasifican dígitos escritos a mano, pero en lugar de combinar ambas técnicas en un solo marco de trabajo, usan los coeficientes de la transformada ondita discreta (*Discrete Wavelet Transform*, DWT, por sus siglas en inglés) como entradas para múltiples CNN. Reemplazar la capa de poleo es la práctica más común al usar onditas dentro de la CNN (como en Ferrá *et al.*, 2019 y Williams y Li, 2018). Estos estudios la utilizan para restaurar imágenes o detectar tumores en el pulmón, entre otras aplicaciones.

Sin embargo, aunque estos últimos modelos presentan un avance con el uso del MRA en el interior

de la CNN, ninguno toma en cuenta las ventajas de los poleos convencionales. Incluso en los resultados obtenidos por Ferrá y colaboradores (2019) se ve que el poleo máximo tiene los mejores resultados. En este sentido, un modelo híbrido que combine las propiedades y ventajas del modelo convencional y del MRA podría tener mejores resultados.

MODELO PROPUESTO

El modelo propuesto por nosotros (Treviño y Alarcón, 2022) es un método de poleo y utiliza dos etapas de procesamiento. En la primera etapa del modelo se realiza un preprocesamiento para mejorar la calidad de las imágenes IR, como se muestra en la Figura 5.

Uno de los primeros retos de este tipo de imágenes es que tienen mucho ruido producto del sentido térmico, además de que suelen ser un poco difusas. Por esta razón, se mejoró la imagen usando una transformada discreta de onditas mediante la cual se aumenta el valor de los detalles para darle más definición y claridad a la misma, además de un filtro con un umbral de ruido para eliminar el ruido inicial. El resultado de estos bloques forma el primer canal de la imagen final. El siguiente bloque produce una imagen binaria que busca resaltar los contornos y los bordes para formar el segundo canal. El tercer canal puede o no existir. En caso de que se desee utilizar, sería la imagen original sin ningún procesamiento. Adicionalmente, se usa una versión simplificada de la CNN para detectar objetos en imágenes IR, lo cual requiere de menos conexiones, capas, operaciones matemáticas y tiempo. Esta CNN cuenta con cuatro capas de convolución, dos de poleo y una de clasificación, además de usar la unidad lineal rectificadora con fuga, como función de activación.

Por otro lado, se implementó el esquema de elevación como una capa embebida más dentro de la CNN configurada para tres onditas llamadas en honor de sus creadores Haar, Daubechies 4 y Daubechies 6. Cada una con características únicas para extraer características específicas de las imágenes, como se muestra en la Figura 6. De esta

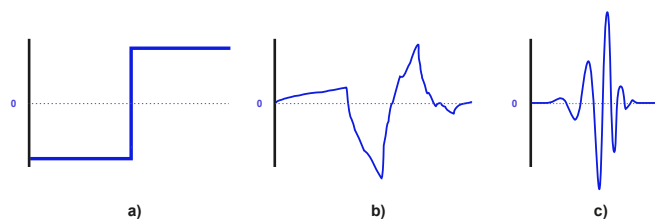


Figura 6. Ejemplos de tres onditas, a) Haar, b) Daubechies 4 y c) Daubechies 6.

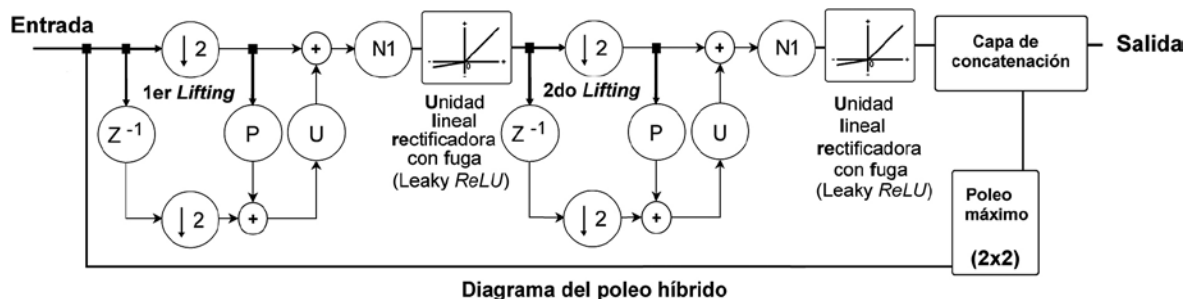


Figura 7. Modelo híbrido propuesto para mejorar la precisión de la CNN.

manera, el modelo híbrido no solo se ha probado con los diferentes poleos convencionales (máximo, mixto, promedio y estocástico), sino con tres onditas (*wavelets*) diferentes. Inicialmente, el modelo de poleo híbrido que combina el método de poleo máximo con la ondita Haar ha mejorado la precisión de la CNN. Otra ventaja del modelo híbrido es que solo requiere los coeficientes de aproximación del bloque de MRA.

Este modelo incorpora en paralelo tanto la técnica del poleo máximo como la del esquema de elevación en 2D para la ondita Haar, como se muestra en la Figura 7. Dado que el modelo está integrado dentro de la CNN, todos sus bloques también son capas de red. El modelo se construye incorporando el proceso de reducción de muestreo ($\downarrow 2$) y los valores de los parámetros P , U y $N1$. Sin embargo, solo las salidas de las aproximaciones están conectadas al siguiente bloque, que es una unidad lineal rectificadora con fuga. Esta capa es una función de activación que mantiene la salida igual a la entrada para los valores positivos y los negativos normalmente los elimina, pero en la configuración con fuga solo los reduce para no darles mucha importancia y, a la vez, evitar que se pierdan. Al mismo tiempo, el bloque de poleo máximo

realiza una reducción de muestreo con una región de poleo de 2×2 , paso de 2×2 sin relleno. Como último paso, la salida producida por el esquema de elevación (que pasa a través de una segunda unidad lineal rectificadora con fuga) y la salida generada por el poleo máximo, se unen mediante una capa de concatenación. La capa de concatenación se usa con frecuencia en las CNN para unir dos o más bloques de información en uno. En este caso, la capa vincula ambos mapas de características para enriquecer el proceso de poleo a pesar de la inherente reducción del proceso. Finalmente, se probaron cuatro poleos convencionales y diversas combinaciones de modelos usando MRA; sin embargo, el modelo con los mejores resultados, como se esperaba, fue el modelo híbrido que usa el poleo máximo en paralelo con el esquema de elevación.

CONCLUSIONES

La CNN es una técnica muy útil para extraer características y clasificar objetos. La precisión con la que actúan estos sistemas es esencial en cualquier aplicación. Por otro lado, a pesar de su uso

Método de Poleo	Precisión
Poleo promedio	98.7 %
Poleo máximo	99.0 %
Poleo mixto	97.9 %
Poleo estocástico	98.4 %
Modelo híbrido: ondita Haar	99.4 %
Modelo híbrido: ondita Daubechies 4	99.4 %
Modelo híbrido: ondita Daubechies 6	99.2 %

Tabla 1. Resultados de los diferentes métodos de poleo

tan generalizado y de la cantidad de aplicaciones desarrolladas, sigue siendo importante lograr un buen desempeño, en especial cuando se usan imágenes con bases de datos limitadas o de difícil acceso para el entrenamiento de las redes, como en el caso de las imágenes IR. Para contribuir a la solución de este problema se propuso mejorar el área de mayor pérdida de información dentro de la CNN. El proceso de reducción del tamaño de la información mientras se extraen características es necesario para que la CNN puede funcionar eficientemente, pero también se sabe que es en las capas de poleo donde se pierden muchos de estos rasgos. Por ello, se incorporó un método que, a diferencia de los poleos convencionales que retienen un valor convencional (valor máximo o promedio), el MRA mantiene las partes esenciales de la información. Estas partes son los cuatro coeficientes. En nuestro caso, solo usamos el coeficiente en que se mantiene la mayor parte de la información. Creemos que para implementar este proceso dentro de la CNN, es necesario un tipo de análisis conocido como el esquema de elevación que puede incorporarse en la arquitectura de la red como una capa más, sustituyendo a las capas de poleo. Los resultados presentados en la Tabla 1 indican la precisión de

la red convolucional clasificando objetos utilizando únicamente como poleo las capas indicadas como método de poleo. Por ejemplo, en el caso del poleo promedio la precisión es de 98.7 %, esto quiere decir que de cada 1,000 objetos introducidos a la red convolucional con esa configuración se clasificaron correctamente 987, mientras que 13 objetos fueron mal clasificados o no pudieron identificarse en ninguna de las categorías utilizadas. Al remplazar cada capa de poleo con un modelo híbrido, como el mostrado en la Figura 7.

La red convolucional presentó un aumento en comparación con los cuatro métodos de poleo convencional. Por ejemplo, entre el modelo híbrido con Haar y el poleo mixto hay una diferencia de 1.5 %, mientras que con el poleo estocástico la diferencia es de 1 %.

Finalmente, también el caso del poleo máximo existe una diferencia, aunque menor, de 0.4 %. También se probaron diversos métodos híbridos con menores resultados, uno de los cuales utilizó el método mixto y su diferencia fue de 2 %. A pesar de que la CNN existe desde hace varias décadas, su uso se hizo posible recientemente gracias al desarrollo del *hardware*. Sin embargo, hablamos de un área que aún tiene muchos temas por investigar, su uso es empírico en cuanto a su diseño (número de neuronas y capas), o simplemente se usa una red ya probada (GoogleNet, VGG, etc.). En cuanto a su aprendizaje, el sobre-entrenamiento sigue siendo un problema porque la red memoriza en lugar de aprender, y aunque hay técnicas no lineales muy efectivas para atacar este inconveniente, aún se requieren respuestas matemáticamente formales.

R E F E R E N C I A S

- Bastidas M, Gruson A, Polanía L, Fujieda S, Prieto F, Takayama K and Hachisuka T (2020). Deep Adaptive Wavelet Network. *Winter Conference on Applications of Computer Vision 2020*:3100-3108.
- Coffey V (2011). Seeing in the Dark: Defense Applications of IR imaging. *Optics and Photonics News* 22:26-31.
- Ferrá A, Aguilar E and Radeva P (2019). Multiple Wavelet Pooling for CNNs. *Lecture Notes in Computer Science* 11132:671-675.
- Larkin P (2018). *IR and Raman Spectroscopy, principles and spectral interpretation*. Holanda: Ámsterdam.



© Enrique Soto. Serie "Mofles", 2010.

Lee C, Gallagher P and Tu Z (2018). Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40:863-875.

Sweldens W (1998). The lifting scheme: A Construction of Second Generation Wavelets. *Siam Journal on Mathematical Analysis* 29:511-546.

Treviño D and Alarcón V (2022). Hybrid Pooling with Wavelets for Convolutional Neural Networks. *Journal of Intelligent & Fuzzy Systems* 42:4327-4336.

Williams T and Li R (2016). Advanced Image Classification using Wavelets and Convolutional Neural Networks. *15th IEEE International Conference on Machine Learning and Applications* 2016:233-239.

Williams T and Li R (2018). Wavelet Pooling for Convolutional Neural Networks. *6th International Conference on Learning Representations*. Canadá: Vancouver.

Yu D, Wang H, Chen P and Wei Z (2014). Mixed Pooling for Convolutional Neural Networks. *Lecture Notes in Computer Science* 8818:364-375.

Zeiler M and Fergus R (2013). Stochastic pooling for regularization of Deep convolutional neural networks. *1st International Conference on Learning Representations*. E.E.U.U.: Scottsdale

G L O S A R I O

Análisis multirresolución (Multi-resolution Analysis, MRA). Se realiza aproximando una función $f(t)$ en distintos niveles de resolución, para producir una descomposición multi-escala.

Eficiencia. Lograr cumplir adecuadamente una tarea o función.

Esquema de elevación (Lifting Scheme). Es un análisis multirresolución de segunda generación. Es una transformada discreta de onditas (wavelet) que está configurada en bloques dependiendo de la función de ondita que se utilice. Los bloques usan operaciones matemáticas simples para lograr descomponer o reconstruir imágenes en sus componentes básicos llamados coeficientes.

Hardware. Son los elementos físicos de una computadora, tanto internos como externos.

Ondita (Wavelet). Es una forma de onda o breve oscilación, que aumenta o disminuye, y que tiene una duración limitada y un valor promedio de cero.

Poleo. Es un componente básico en una CNN que consolida las características aprendidas. Se usa para reducir la dimensionalidad de los datos, así como de parámetros requeridos por la CNN.

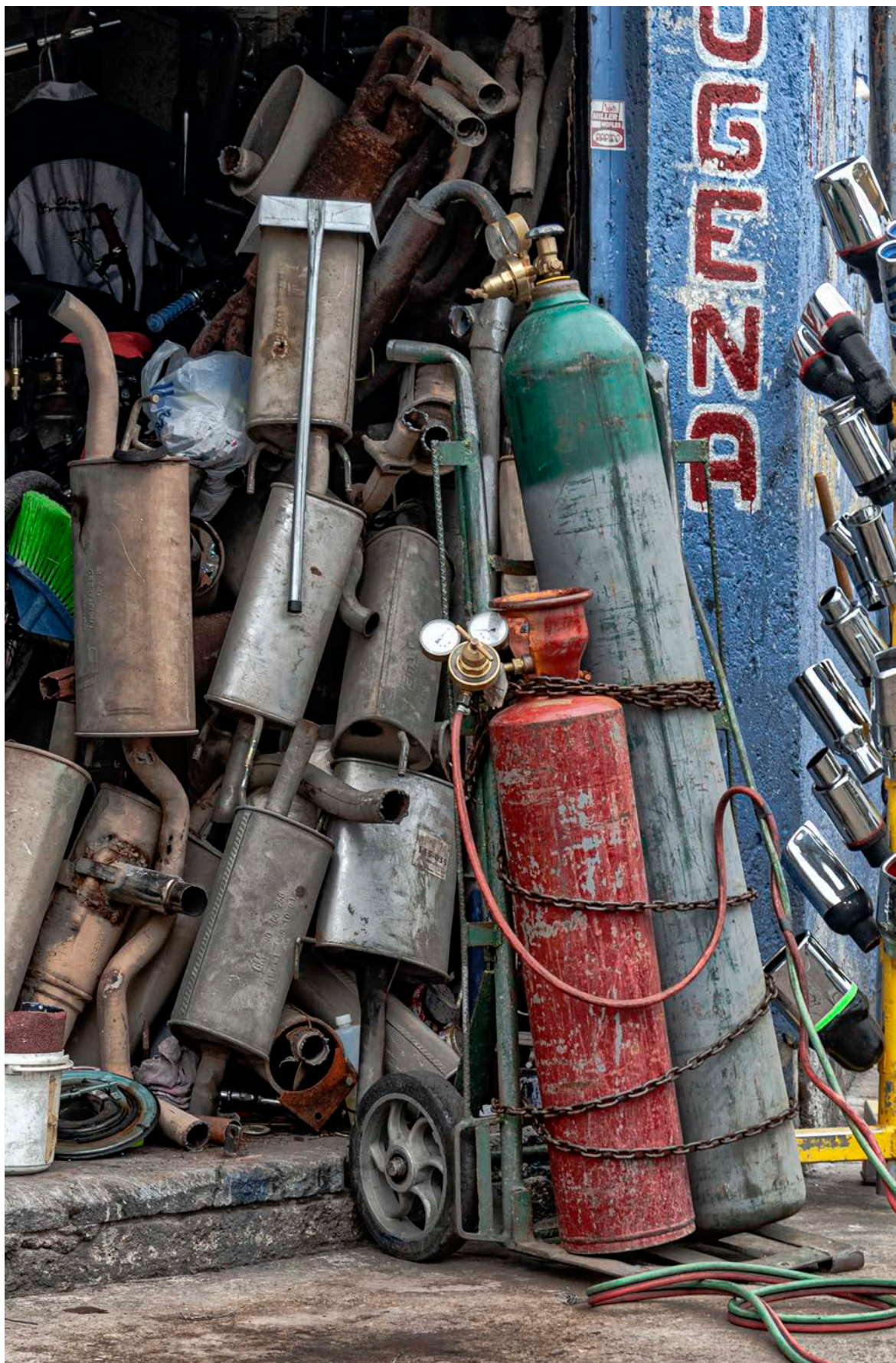
Precisión. La precisión se refiere a la cercanía de los valores obtenidos aplicando los mismos parámetros.

Red neuronal convolucional (Convolutional Neural Network, CNN). Es una arquitectura de red que aprende directamente de los datos. Sus neuronas son nodos que realizan operaciones matemáticas simples para actualizar sus variables llamadas pesos, con los que logran identificar patrones que la red pueda usar para reconocer objetos en una imagen.

Sobreentrenamiento (Overfitting). Es un comportamiento no deseado en el aprendizaje automático. Durante el entrenamiento el sistema empieza a memorizar los datos aparentando mejorar su desempeño; sin embargo, su desempeño es deficiente cuando se utilizan datos nuevos.

Transformada discreta de onditas (Discrete Wavelet Transform, DWT). Es una transformada que utiliza onditas muestreadas discretamente. Tiene la ventaja que su resolución es temporal, es decir que muestra tanto la frecuencia como su ubicación en el tiempo.

Daniel Treviño Sánchez
Vicente Alarcón Aquino
Universidad de las Américas Puebla
Departamento de Computación
Electrónica y Mecatrónica
vicente.alarcon@udlap.mx



© Enrique Soto. Serie "Mofles", 2006