

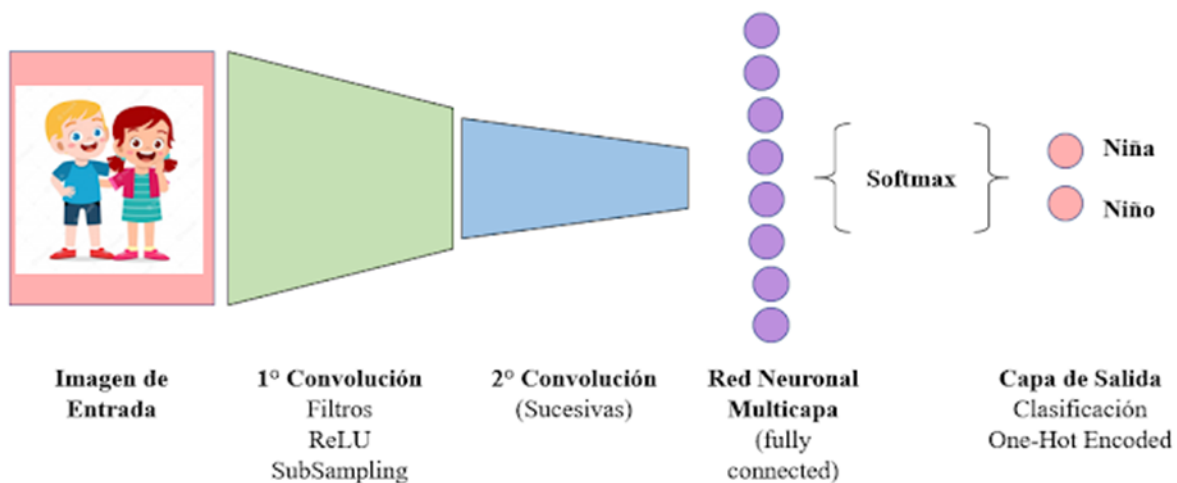


REDES NEURONALES CONVOLUCIONALES

Definición

A diferencia de las redes neuronales, las redes neuronales convolucionales son un subconjunto de aprendizaje automático y están en el centro de los algoritmos de aprendizaje profundo (Deep learning). 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.

Arquitectura de Redes Neuronales Convolucionales – CNN





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.

Funcionamiento

Las CNN se distinguen de otras redes neuronales por su mejor desempeño con entradas de señal de imagen, voz o audio. Tienen tres tipos principales de capas, que son:

- Capa convolucional
- Capa de agrupamiento
- Capa totalmente conectada (FC)

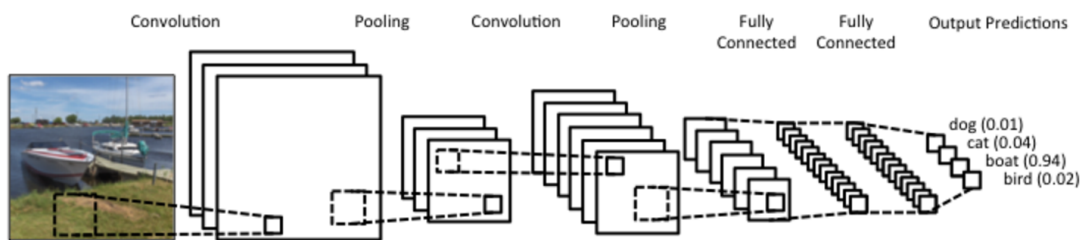
La capa convolucional es la primera capa de una red convolucional. Si bien a las capas convolucionales las pueden seguir capas convolucionales adicionales o capas agrupadas, la capa totalmente conectada es la capa final. Con cada capa, la CNN aumenta su complejidad, identificando mayores porciones de la imagen. Las primeras capas se enfocan en características simples, como colores y bordes. A medida que los datos de la imagen avanzan a través de las capas de CNN, se comienzan a reconocer elementos o formas más grandes del objeto, hasta que finalmente se identifica el objeto previsto.

La capa convolucional es el pilar central de una CNN, y es donde ocurre la mayor parte del cálculo. Requiere algunos componentes, que son datos de entrada, un filtro y un mapa de características. Supongamos que la entrada será una imagen en color, que se compone de una matriz de píxeles en 3D. Esto significa que la entrada tendrá tres dimensiones (altura, ancho y profundidad) que corresponden a RGB en una imagen.



También hay un detector de características, conocido como kernel o filtro, que se moverá a través de los campos receptivos de la imagen, verificando si la característica está presente. Este proceso se conoce como convolución.

El detector de características es una matriz bidimensional (2-D) de pesos, que representa parte de la imagen. Aunque esto puede variar, el tamaño del filtro suele ser una matriz 3 x 3; esto también determina el tamaño del campo receptivo. Luego, el filtro se aplica a un área de la imagen y se calcula un producto de punto entre los píxeles de entrada y el filtro. Este producto de punto se introduce entonces en una matriz de salida. Después, el filtro cambia a pasos agigantados, repitiendo el proceso hasta que el kernel ha barrido toda la imagen. El resultado final de la serie de productos de puntos de la entrada y el filtro se conoce como mapa de características, mapa de activación o característica convolucionada.



Después de cada operación de convolución, una CNN aplica una transformación de Unidad lineal rectificada (ReLU) al mapa de características, introduciendo la no linealidad en el modelo.

Como se mencionó anteriormente, otra capa de convolución puede seguir a la capa de convolución inicial. Cuando esto sucede, la estructura de la CNN puede convertirse en jerárquica, ya que las capas posteriores pueden ver los píxeles dentro de los campos receptivos de las capas anteriores. Por ejemplo, supongamos que intentamos determinar si una imagen contiene una bicicleta. Pensemos en la bicicleta como una suma de partes.



Está compuesta por una estructura, un manubrio, llantas, pedales, etc. Cada parte individual de la bicicleta constituye un patrón de nivel inferior en la red neuronal, y la combinación de sus piezas representa un patrón de nivel superior, creando una jerarquía de características dentro de la CNN.

Aplicaciones

Las CNN forman parte del machine learning, una rama de la Inteligencia Artificial que trabaja para que las máquinas aprendan de forma autónoma. Principalmente, se usan para identificar patrones de conjuntos de datos, por ejemplo, el reconocimiento de voz, audio o clasificar diferentes categorías de imágenes, de tal forma que permiten la detección de objetos, para después servir como base para la toma de decisiones.

Además, son muy eficaces para el procesamiento del lenguaje humano. Se pueden utilizar para la traducción automática, las redes neuronales aprenden a traducir a un determinado idioma gracias al análisis de patrones del lenguaje.

Al mismo tiempo, pueden predecir la toma de decisión. En este sentido, son muy favorecedoras para analizar el comportamiento de los clientes sobre las compras en un establecimiento, y predecir los posibles riesgos de enfermedades de los pacientes.

Ventajas y desventajas

Entre los principales beneficios destaca su eficacia para la identificación de patrones visuales. Por este motivo, se usan para clasificar las imágenes, que anteriormente podían resultar complejas ante redes convolucionales de estructura convencional.



Asimismo, cuentan con una alta capacidad para procesar grandes volúmenes de datos visuales. Esta tecnología aprende de su experiencia, por tanto, mejora sus procedimientos con el paso del tiempo.

No obstante, estas redes no son perfectas, también presentan una serie de debilidades. En primer lugar, pueden necesitar grandes cantidades de datos para poder operar correctamente. Además, suelen precisar mucho tiempo de instrucción para su adecuada optimización, y además necesitan equipos más potentes.