
REDES NEURONALES CONVOLUCIONALES EN LA CLASIFICACIÓN DE CULTIVOS

5 de agosto de 2019

Fernando Fortanel Rojas
Universidad Nacional Autónoma de México
Unidad Cuernavaca del Instituto de Matemáticas

Índice general

1. Introducción	2
1.1. Antecedentes	2
2. Objetivos e Hipótesis	4
3. Metodología	6
3.1. Redes neuronales	6
3.1.1. Perceptrón	6
3.2. Redes neuronales convolucionales	7
3.3. Convolucion de matrices	8
3.4. Filtrado de imagenes	9
3.5. Estructura de una Red Neuronal Convolucional	10
3.6. La capa convolucional	11

Capítulo 1

Introducción

Debido a la gran cantidad de imágenes satelitales disponibles, se han desarrollado múltiples tecnologías para el aprovechamiento de esta información en el campo de los sensores remotos. Un ejemplo es el de la clasificación automática de cultivos, la cual tiene particular relevancia ya que es útil para la estimación del rendimiento, la evaluación económica, la transportación de los cultivos, así como la estimación de los recursos hídricos necesarios para el cultivo. Distintas técnicas clásicas de clasificación se han usado para el desarrollo de esta tecnología como son, máquinas de soporte vectorial (SVM), k vecinos más cercanos (KNN), clasificación de máxima verosimilitud (MLC), entre otras técnicas. Sin embargo, han sido ineficientes bajo ciertas circunstancias de la vegetación, del clima y de la geografía del área a ser analizada.

Los métodos de clasificación con sensores remotos consiste de dos partes principales, la primera un extractor de características, la cual transforma datos espaciales, espectrales y/o temporales en vectores de características discriminativas y en un clasificador que etiqueta a cada vector de características en una cierta clase. En este sentido es donde las redes neuronales convolucionales podrían lograr un buen desempeño dado la capacidad de estas para detectar características, pero antes de hablar más afondo de esto, primero empezaremos describiendo algunos resultados previos.

1.1. ANTECEDENTES

En 2016, Hou, Wang, Xu, Reid, y Han [5] mostraron que distintos tipos de superficie presenta un comportamiento espectral distinto, es decir la vegetación verde o saludable, la vegetación seca, el suelo desnudo y otras superficies, son distintas y presentan variacio-

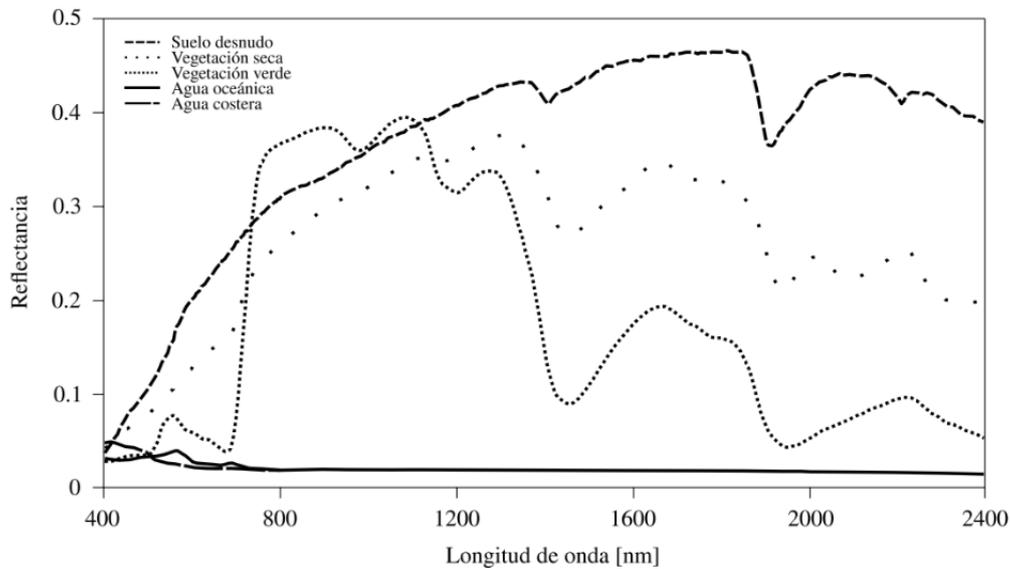


Figura 1.1: La reflectancia de distintas superficies (tomado de [6])

nes en su reflectancia para distintas zonas del espectro. Esto nos permite conjeturar, la posibilidad de identificar que tipo de superficie es, tan solo teniendo la información de la intensidad de reflectancia que presenta un objeto, como se muestra en la figura 1. Además es clara la información que aporta las imágenes multitemporales de un cultivo, dado que pueden mostrar los distintos cambios en el cultivo, por ejemplo el ciclo del arroz incluye las siguientes etapas germinación, macollaje, diferenciación de espigas, rumbo y floración, ordeño y una etapa madura. También es claro que esta información multitemporal es útil para discriminar los cultivos de los objetos estáticos con respecto al tiempo, por ejemplo construcciones o elementos geográficos.

En años recientes, distintos estudios relacionados a las redes neuronales han obtenido buenos resultados, sin embargo una red neuronal convolucional tradicional dos dimensional(2D), principalmente diseñada para imágenes a color (RGB), falla en la habilidad para extraer información de características extras. Para superar esto Kussul([1]) introdujo dos CNN para la clasificación de cultivos, una CNN 2D para aprender características espaciales y otra uno-dimensional (1D) CNN para aprender características espectrales, sin embargo, esta estrategia requiere procesamiento empírico adicional para combinar las características aprendidas.

Capítulo 2

Objetivos e Hipótesis

El objetivo de este trabajo es utilizar métodos matemáticos y de aprendizaje profundo para clasificar cultivos, a partir de imágenes satelitales de diferentes regiones del país, en especial se investigara y analizara el desempeño obtenido mediante un tipo especial de redes neuronales convolucionales en la clasificación de cultivos, junto con información multiespectral, por lo cual, se investigara todo lo referente a las redes convolucionales tres-dimensionales (3D) y como estas podrían lograr mejores resultados en la clasificación de cultivos, al agregar nuevas características tanto espectrales como temporales, no consideradas con CNN tradicionales.

Por ultimo se analizara una técnica propuesta en el trabajo de Shunping Ji, Chi Zhang, Anjian Xu, Yun Shi y Yulin Duan el cual es una estrategia de aprendizaje semi-supervisado y semi-automático para reducir el numero de datos, necesarios para lograr el entrenamiento del modelo. El motivo para reducir esta información es debido a la dificultad de obtenerlos, ya que esto depende de que existan condiciones climáticas óptimas.

Por lo tanto los objetivos específicos son:

- 1.- A través de técnicas de machine learning y CNN, llevar a cabo una clasificación de cultivos en ciertas regiones del país.
- 2.- Construir indicadores relativos a los diferentes tipos de cultivos, superficies, rendimientos, etc.
- 3.- Comparar los resultados entre las redes neuronales convolucionales tradicionales (2D) con respecto las tres-dimensionales (3D).

- 4.- Implementar la estrategia de aprendizaje semi-supervisada y semi-automático, y comparar el desempeño al usar y no usarla.

Hipótesis

Dado la información brinda anteriormente en la introducción, la hipótesis de este trabajo es que es posible usar métodos de aprendizaje de máquina y redes neuronales para reconocer exitosamente distintos tipos de cultivos.

Capítulo 3

Metodología

3.1. REDES NEURONALES

Una red neuronal artificial (ANN) es un modelo matemático que esta inspirados en la forma en como trabajan las neuronas del cerebro humano. Actualmente es una de las técnicas matemáticas más utilizadas en el machine learning.

3.1.1. Perceptrón

La unidad básica de las redes neuronales, es el perceptrón, este recibe los datos de otros perceptrones o de una fuente externa, despues procesa la informaciñ y regresa el resultado. Cada entrada tiene asociado un peso (w), el cual es ponderado con el dato de entrada ,además el perceptrón cuenta con una función de activación, la cual es la que da el resultado, el cual es evaluar la función en los datos ponderados con los pesos.

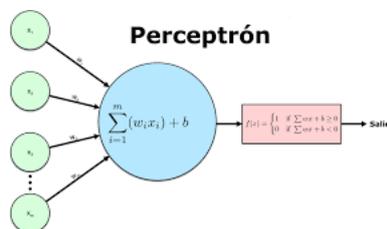


Figura 3.1: Unidad básica de una red neuronal

Algunas de las funciones de activación más populares son la sigmoide, la tanh y ReLU. El objetivo de la función de activación es poder separar los datos, donde algunos de los valores "activan" la neurona o no, así poder separar la información en dos categorías, por

cada perceptrón. Usualmente se pide que esta función sea derivable en todo su dominio, sin embargo como la función ReLU muestra, esto no siempre ocurre.

Para lograr que la red neuronal aprenda, se modifican los pesos para lograr que el perceptrón se active o no. Para lograr esto se tienen múltiples algoritmos, entre ellos el más famoso es el de descenso por gradiente, y este es el motivo por el se trata que la función de activación sea derivable en todo punto. Con esta unidad básica que es el perceptrón, la red

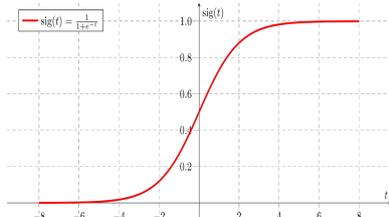


Figura 3.2: Sigmoide

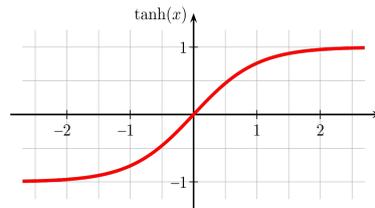


Figura 3.3: Tanh

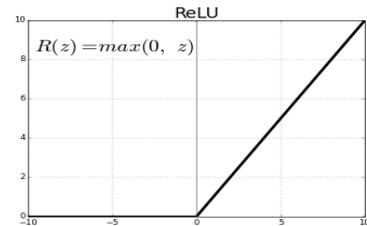


Figura 3.4: ReLU

neuronal se compone de capas, donde cada capa consiste de un conjunto de perceptrones, donde los datos entrantes son obtenidos de una fuente externa de la capa. Debe notarse que los perceptrones no necesariamente tiene la misma función de activación. Estas capas se concatenan para formar a la red neuronal.

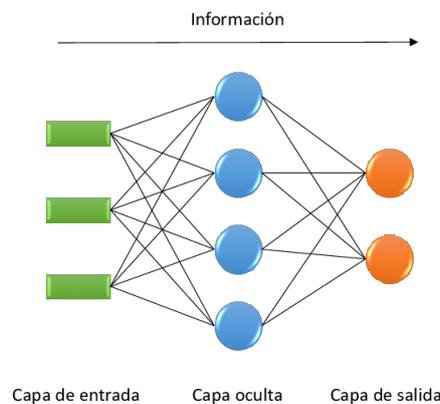


Figura 3.5: Red neuronal artificial

3.2. REDES NEURONALES CONVOLUCIONALES

Las redes neuronales convolucionales o CNN por sus siglas en ingles, son un tipo de redes neuronales usualmente utilizadas en el procesamiento de imágenes y sonidos. Este tipo de redes ha tenido un gran éxito en múltiples aplicaciones debido principalmente a

su capacidad de reconocimiento de características. El nombre de este tipo de redes hace referencia a una operación matemática llamada convolución, la cual tiene diversas aplicaciones en la probabilidad, estadística, visión computacional, procesamiento natural de lenguaje.

Entonces las redes neuronales convolucionales son redes neuronales simples que usan la operación de convolución en vez de la multiplicación de matrices en al menos una de sus capas.

3.3. CONVOLUCION DE MATRICES

En la forma más general, la convolución es una operación matemática que toma dos funciones y genera una nueva función, la cual expresa como la forma de una es modificada por la segunda. La primer función usualmente es la función sobre la cual se ejerce el cambio mientras que la segunda función es la que modifica a la primera, usualmente se les llama canal y kernel respectivamente.

Normalmente en el machine learning, tanto el canal como el kernel son matrices, por lo que la convolución es una convolución de matrices.

Definición Dada una matriz $A_{m \times n}$ y una matriz $C_{(2N+1) \times (2N+1)}$ con $2N+1 < m, n$ se define la convolución de las matrices A y C como una nueva matriz $D = A * C$ definida a partir de la expresión

$$d_{ij} = \frac{1}{c} \sum_{r=1}^{2N+1} \sum_{s=1}^{2N+1} a_{i-N+r-1, i-N+r-1} c_{r, s} \quad (3.1)$$

donde $c = \sum_{i,j=1}^{2N+1} c_{i,j}$ (si $c = 0$ se toma $c = 1$). Obsérvese que $d_{i,j}$ sólo está definido para $i = N+1, \dots, m-N-1$ y $j = N+1, \dots, n-N-1$. Donde la matriz A es el canal y la matriz C es el kernel.

Las dos siguientes figuras muestran el proceso de convolución de una matriz dada por otra de orden 3×3 .

$$\begin{array}{ccccc}
 35 & 40 & 41 & 45 & 50 \\
 40 & 40 & 42 & 46 & 52 \\
 42 & 46 & 50 & 55 & 55 \\
 48 & 52 & 56 & 58 & 60 \\
 56 & 60 & 65 & 70 & 75
 \end{array}
 *
 \begin{array}{ccc}
 -2 & -1 & 0 \\
 -1 & 1 & 1 \\
 0 & 1 & 2
 \end{array}
 =
 \begin{array}{ccccc}
 & & & & \\
 & & 78 & & \\
 & & & & \\
 & & & & \\
 & & & &
 \end{array}$$

$$35 \cdot (-2) + 40 \cdot (-1) + 41 \cdot 0 + 40 \cdot (-1) + 40 \cdot 1 + 42 \cdot 1 + 42 \cdot 0 + 46 \cdot 1 + 50 \cdot 2 = 78$$

$$\begin{array}{ccccc}
 35 & 40 & 41 & 45 & 50 \\
 40 & 40 & 42 & 46 & 52 \\
 42 & 46 & 50 & 55 & 55 \\
 48 & 52 & 56 & 58 & 60 \\
 56 & 60 & 65 & 70 & 75
 \end{array}
 *
 \begin{array}{ccc}
 -2 & -1 & 0 \\
 -1 & 1 & 1 \\
 0 & 1 & 2
 \end{array}
 =
 \begin{array}{ccccc}
 & & & & \\
 & & & 87 & \\
 & & & & \\
 & & & & \\
 & & & &
 \end{array}$$

$$40 \cdot (-2) + 41 \cdot (-1) + 45 \cdot 0 + 40 \cdot (-1) + 42 \cdot 1 + 46 \cdot 1 + 46 \cdot 0 + 50 \cdot 1 + 55 \cdot 2 = 87$$

Figura 3.6: El resultado de la convolución correspondiente a las entradas(2,2) y (3,2)

$$\begin{array}{ccccc}
 35 & 40 & 41 & 45 & 50 \\
 40 & 40 & 42 & 46 & 52 \\
 42 & 46 & 50 & 55 & 55 \\
 48 & 52 & 56 & 58 & 60 \\
 56 & 60 & 65 & 70 & 75
 \end{array}
 *
 \begin{array}{ccc}
 -2 & -1 & 0 \\
 -1 & 1 & 1 \\
 0 & 1 & 2
 \end{array}
 =
 \begin{array}{ccccc}
 35 & 40 & 41 & 45 & 50 \\
 40 & 78 & 87 & 94 & 52 \\
 42 & 98 & 283 & 108 & 55 \\
 48 & 120 & 125 & 127 & 60 \\
 56 & 60 & 65 & 70 & 75
 \end{array}$$

Figura 3.7: Resultado final de la convolución, donde los valores de las esquinas es el valor original.

3.4. FILTRADO DE IMAGENES

Las imágenes digitales pueden ser representadas a través de tres matrices de números en caso de imágenes a color (RGB) y con solo una en caso de una imagen a escala de grises, donde cada número se encuentra entre 0 y 255.

Por lo tanto es posible usar la convolución para modificar estas imágenes digitales, a esta proceso se le conoce como filtrado de imágenes. A continuación se muestra algunos resultados al aplicarle la convolución a una imagen con distintos kernels.

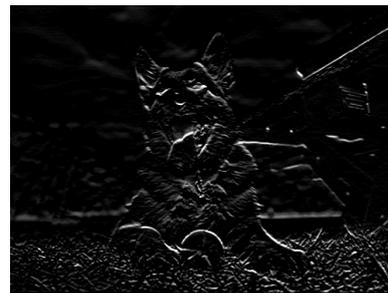


Figura 3.8: Imagen original **Figura 3.9:** Filtro emboss **Figura 3.10:** Filtro sobel

3.5. ESTRUCTURA DE UNA RED NEURONAL CONVOLUCIONAL

Como se menciona anteriormente una red neuronal convolucional es una red neuronal con al menos una capa convolucional, por lo que solo esta capa se describirá a profundidad. Las redes convolucionales se dividen en dos partes principales:

- La parte de extracción de características, en esta parte, la red realizará una serie de convoluciones y de operaciones de pooling, aquí es donde las características son detectadas.
- La parte clásica, aquí las capas de fully connected servirán como un clasificador de las características detectadas.

A continuación se coloca una imagen que ejemplifica las dos partes de las redes convolucionales.

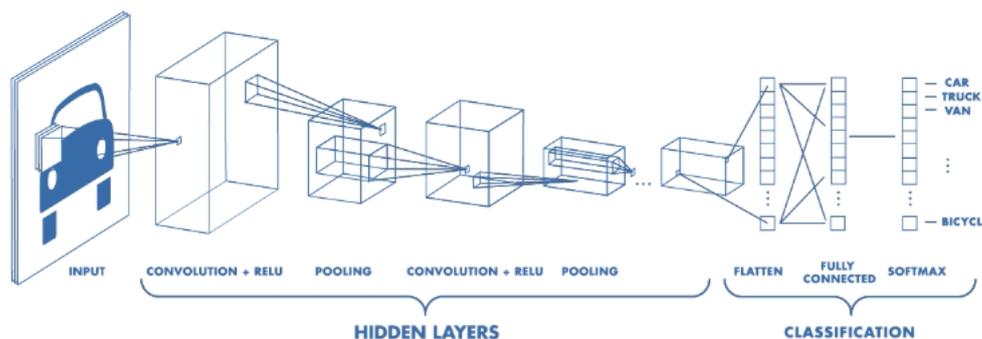


Figura 3.11: La estructura de una red neuronal convolucional estándar

3.6. LA CAPA CONVOLUCIONAL

La capa convolucional se compone de tres partes principales

- La entrada, la cual se compone de múltiples canales, los cuales pueden ser pensados como las imágenes, sin embargo entran en forma de matrices.
- El kernel, el cual es el filtro el cual va ser aplicado la convolución sobre los canales de entrada.
- La salida, la cual consiste del resultado de aplicar la convolución del kernel sobre los canales de entrada después de aplicarle la función de activación, la cual suele ser ReLU (unidad lineal rectificadora por sus siglas en inglés) en el caso de las CNN.

Además, en esta parte se suele aplicar una operación extra que es la de pooling, la cual consiste en reducir la dimensión de los canales resultantes, esto para resumir estadísticamente los valores cercanos y de esta manera mejorar el tiempo de entrenamiento de la red, ya que esta tiene menos datos que tratar. Algunas de las funciones de pooling más usuales son:

- Max-Pooling, la cual consiste de tomar el canal original y dividirla en submatrices cuadradas y tomar el valor máximo de cada submatriz, para generar una nueva submatriz.
- Average-Pooling, esta de igual forma divide el canal original en submatrices y el valor final de cada una de las entradas es el promedio de su correspondiente submatriz.

En la **Figura 3.12** se puede ver un ejemplo de estas dos funciones.

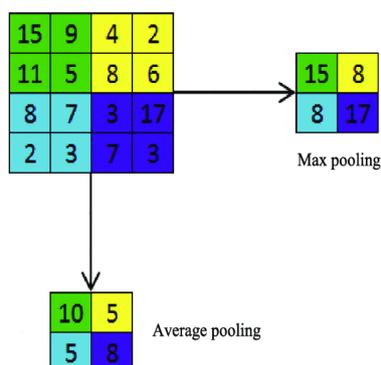


Figura 3.12: Se muestra el proceso de un max-pooling de 2x2 y de average-pooling de 2x2

Bibliografía

- [1] SHUNPING JI, CHI ZHANG , ANJIAN XU, YUN SHI , YULIN DUAN, *3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images. MDPI, 1, 17*, De remote sensing Base de datos , 7 de Enero del 2018
- [2] YAPING CAI, KAIYU GUAN, JIAN PENG SHAOWEN WANG , CHRISTOPHER SEIFERT , BRIAN WARDLOW , ZHAN LI *A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach.*, De Elsevier Base de datos,16 de Marzo de 2018
- [3] IAN GOODFELLOW, YOSHUA BENGIO, AARON COURVILLE., *Deep Learning. Estados Unidos*, MIT Press, 2016.
- [4] WALDO OJEDA BUSTAMANTE, JORGE FLORES VELÁZQUEZ Y RONALD ERNESTO ONTIVEROS CAPURATA., *Uso y manejo de drones con aplicaciones al sector hídrico.*, IMTA, 2018
- [5] HOU, W., WANG, J., XU, X., REID, J. S., Y HAN, D., *An algorithm for hyperspectral remote sensing of aerosols* Journal of Quantitative Spectroscopy and Radiative Transfer, 178, 400-415, 2016.
- [6] BRAULIO ANTONIO VÁZQUEZ RODRÍGUEZ, *ESTIMACIÓN DE LA EVAPOTRANSPIRACIÓN DE CULTIVO DE MAÍZ BAJO RIEGO MEDIANTE PERCEPCIÓN REMOTA*, IMTA 2019.