

Tarea 2: ConvNets (Parte teórica)

Entrega: Jueves 12 de septiembre, 23:59 (Gradescope)

Profesor: Pablo Estévez V.
Auxiliar: Ignacio Reyes J.
Semestre: Primavera 2019

1. Pregunta 1

1. ¿Qué sentido tiene utilizar capas convolucionales con filtros de 1 pixel de alto por 1 pixel de ancho?. Imagine por ejemplo una capa convolucional con 32 filtros de 1 pixel de alto y 1 pixel de ancho, la cual opera sobre imágenes a color (RGB). ¿Qué clase de características podría construir dicha capa convolucional?

Sol:

El filtro 1x1 proporciona una forma de resumir útilmente los mapas de características de entrada. El uso de múltiples filtros 1x1, a su vez, permite la sintonización de la cantidad de mapas de características de entrada para crear, lo que permite aumentar o disminuir la profundidad de los mapas de características según sea necesario.

El filtro 1x1 se puede usar para crear una proyección lineal de una pila de mapas de características.

En el caso de una capa convolucional con 32 filtros de 1x1 pixels, después de la convolución con esta capa obtenemos una capa de 32 de profundidad con el mismo ancho y alto, obteniendo 32 mapas de características.

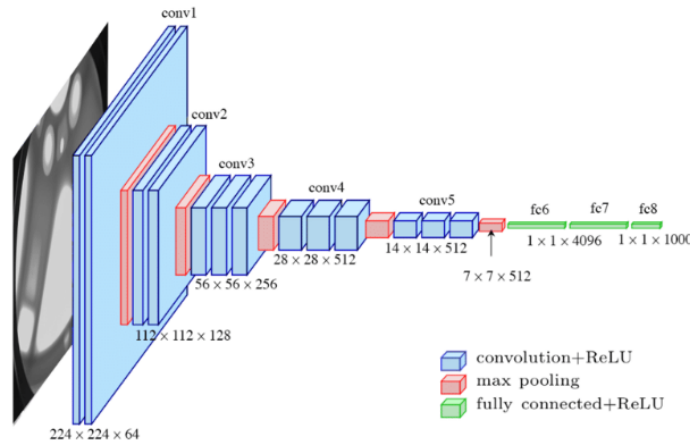
2. Se dice que las redes convolucionales son capaces de aprender características (features) a partir de los datos. ¿A qué se refiere esta afirmación? ¿Cómo se compara con el paradigma de feature engineering?

Sol:

Las redes convolucionales fuerzan la extracción de características locales (que toman ventaja de la correlación) mediante la restricción de los campos receptivos de las unidades ocultas a ser locales.

El paradigma de feature engineering se refiere a dos modelos para reconocer patrones. En el modelo tradicional de reconocimiento de patrones se tiene dos módulos. El módulo de extracción de características, que es hecho a mano, y el clasificador que es automático. Pero en el modelo de redes convolucionales no es necesario, ni razonable hacer los filtros a mano. la red se encarga de ellos, mediante convolución con filtros y backpropagation.

2. Pregunta 2 - Tamaños y formas de los parámetros y activa- ciones



Sol:

Image(RGB): [224x224x3] weights: 0

CONV3-64: [224x224x64] weights: $(3*3*3)*64 = 1,728$

CONV3-64: [224x224x64] weights: $(3*3*64)*64 = 36,864$

maxpool2: [112x112x64] weights: 0

CONV3-128: [112x112x128] weights: $(3*3*64)*128 = 73,728$

CONV3-128: [112x112x128] weights: $(3*3*128)*128 = 147,456$

maxpool2: [56x56x128] weights: 0

CONV3-256: [56x56x256] weights: $(3*3*128)*256 = 294,912$

CONV3-256: [56x56x256] weights: $(3*3*256)*256 = 589,824$

CONV3-256: [56x56x256] weights: $(3*3*256)*256 = 589,824$

maxpool2: [28x28x256] weights: 0

CONV3-512: [28x28x512] weights: $(3*3*256)*512 = 1,179,648$

CONV3-512: [28x28x512] weights: $(3*3*512)*512 = 2,359,296$

CONV3-512: [28x28x512] weights: $(3*3*512)*512 = 2,359,296$

maxpool2: [14x14x512] weights: 0

CONV3-512: [14x14x512] weights: $(3*3*512)*512 = 2,359,296$

CONV3-512: [14x14x512] weights: $(3*3*512)*512 = 2,359,296$

CONV3-512: [14x14x512] weights: $(3*3*512)*512 = 2,359,296$

maxpool2: [7x7x512] weights: 0

FC-4096: [1x1x4096] weights: $7*7*512*4096 = 102,760,448$

FC-4096: [1x1x4096] weights: $4096*4096 = 16,777,216$

FC-1000: [1x1x1000] weights: $4096*1000 = 4,096,000$

Número total de parámetros o pesos = $138344128 \approx 138\text{M}$ parámetros.