

Convolutional neural networks

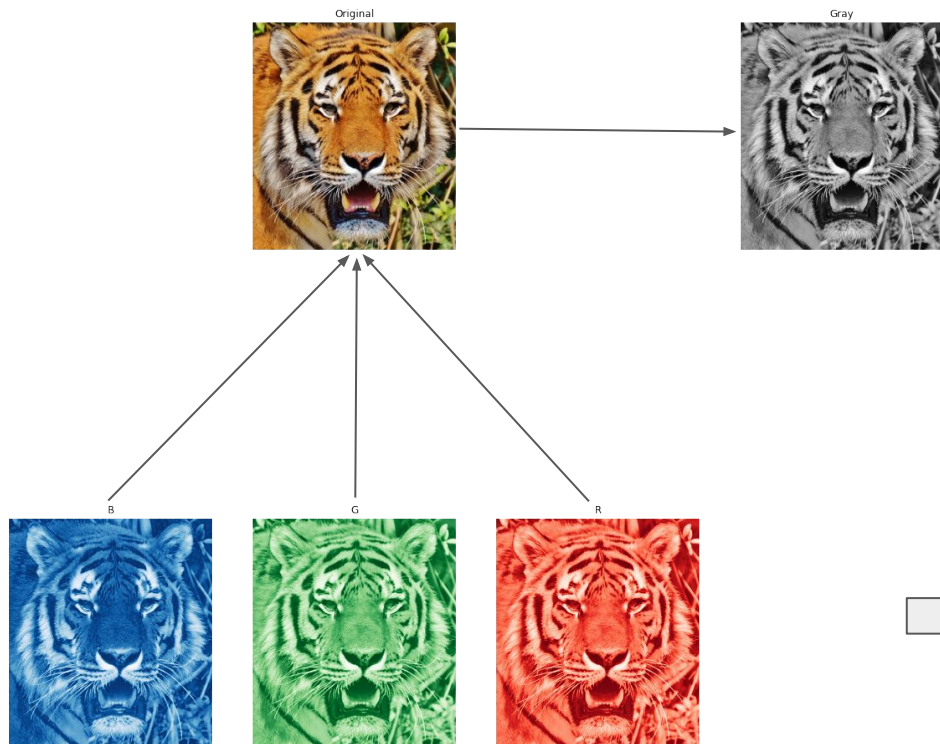
FIC - Ta Dang Khoa

Nội dung

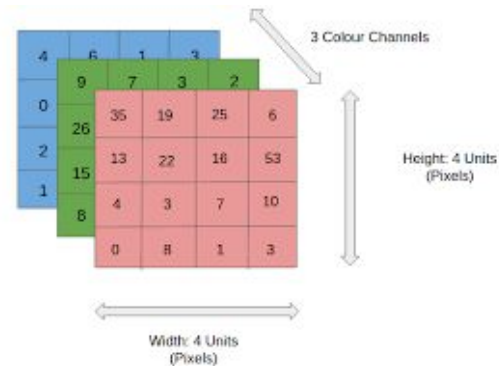
- Giới thiệu
 - Dữ liệu ảnh
 - NN với bài toán phân loại ảnh
- Phép tích chập (convolution)
- Mạng CNN phổ thông
 - Tóm tắt
 - Convolutional layer
 - Activation layer (ReLU)
 - Pooling layer
 - Fully connected layer

Giới thiệu

Dữ liệu ảnh

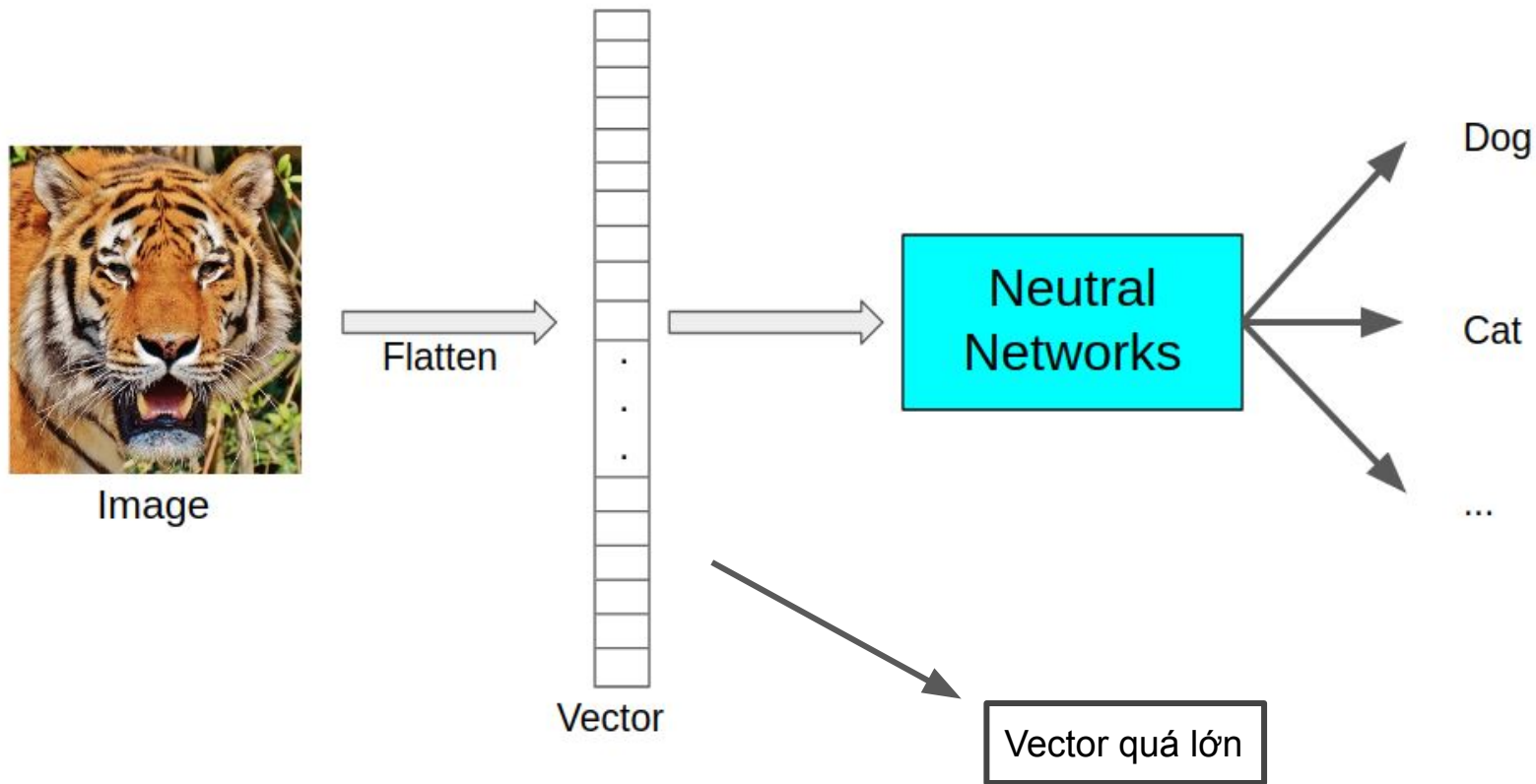


255	184	178	84	129
84	255	255	130	84
78	255	0	0	78
84	130	255	130	84



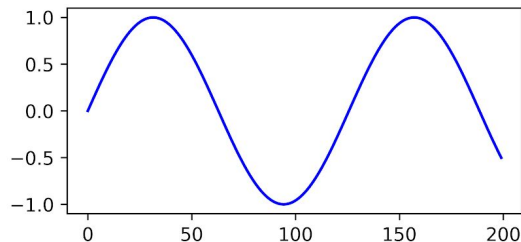
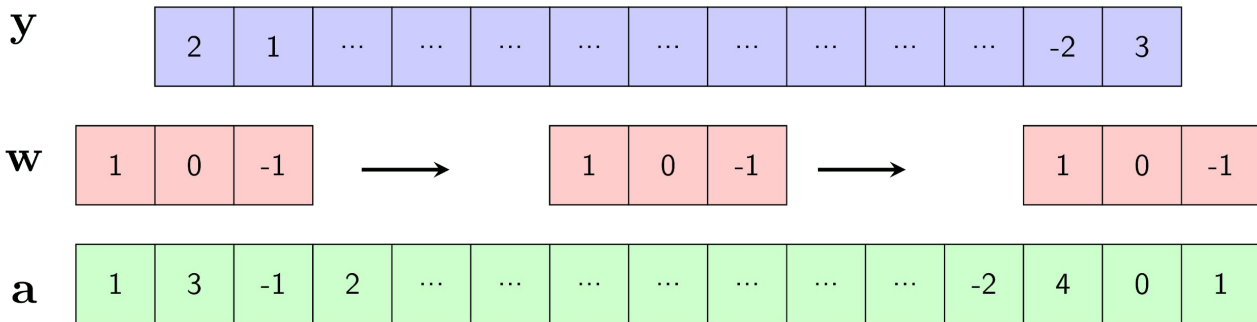
Giới thiệu

NN với bài toán phân loại ảnh

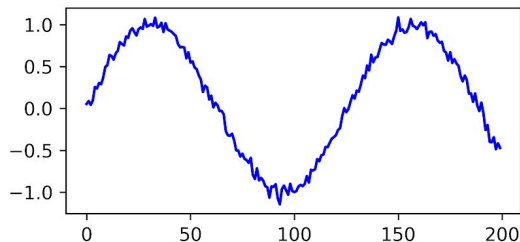


Phép tích chập

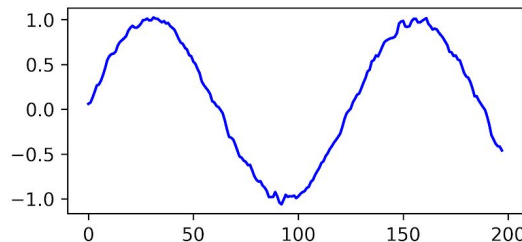
Tích chập một chiều



(a) clean signal



(b) noisy signal



(c) denoised signal

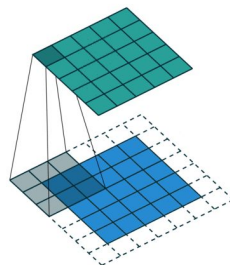
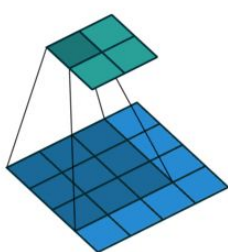
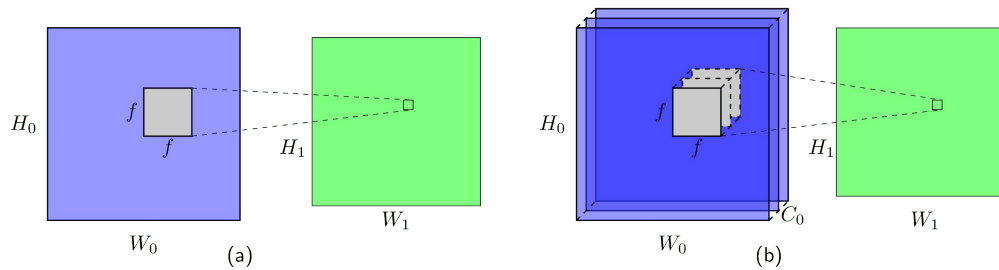
Ví dụ về tích chập một chiều (Nguồn machinelearningcoban.com)

Phép tích chập

Tích chập hai chiều

$$H_1 = \frac{H_0 + 2P - f}{S} + 1$$

$$W_1 = \frac{W_0 + 2P - f}{S} + 1$$



Phép tích chập

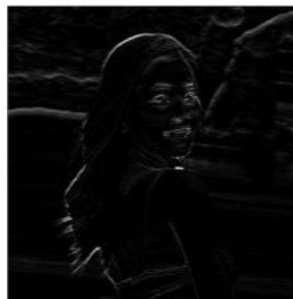
Tích chập hai chiều - Ứng dụng



(a)



(b)



(c)

Hình 6. (a) Ảnh xám gốc. (b) Dò tìm các cạnh thẳng đứng trong ảnh. (c) Dò tìm các cạnh nằm ngang.

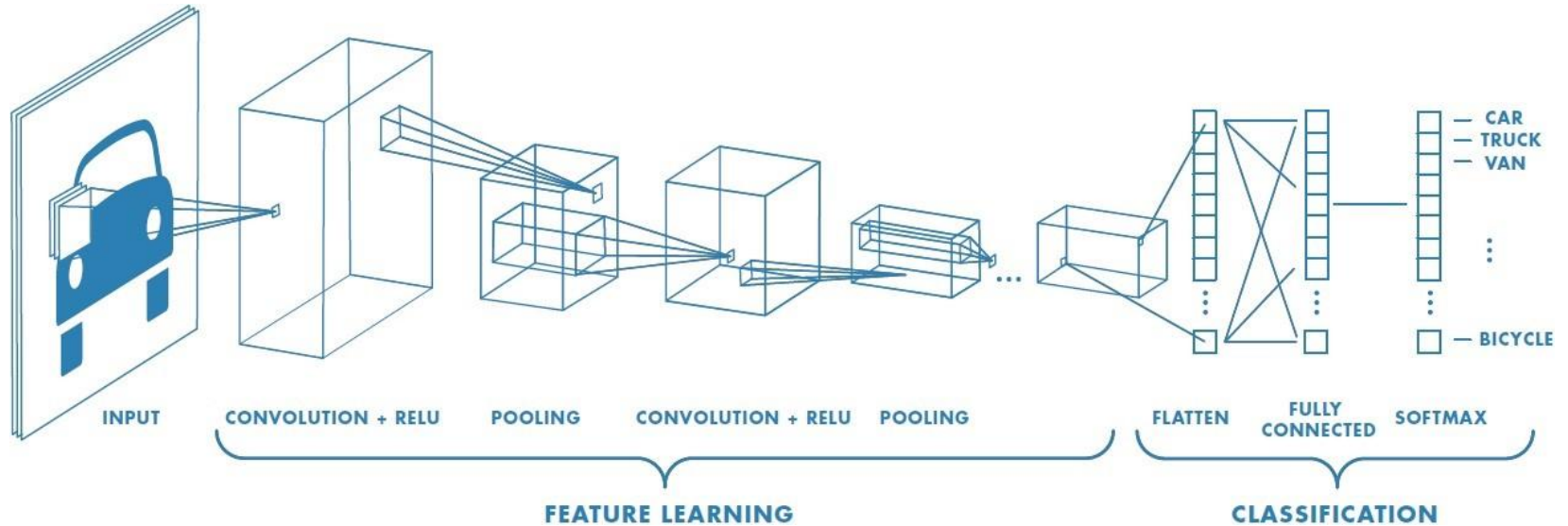
$$w = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

Ví dụ ứng dụng dò cạnh của tích chập

Nguồn machinelearningcoban.com

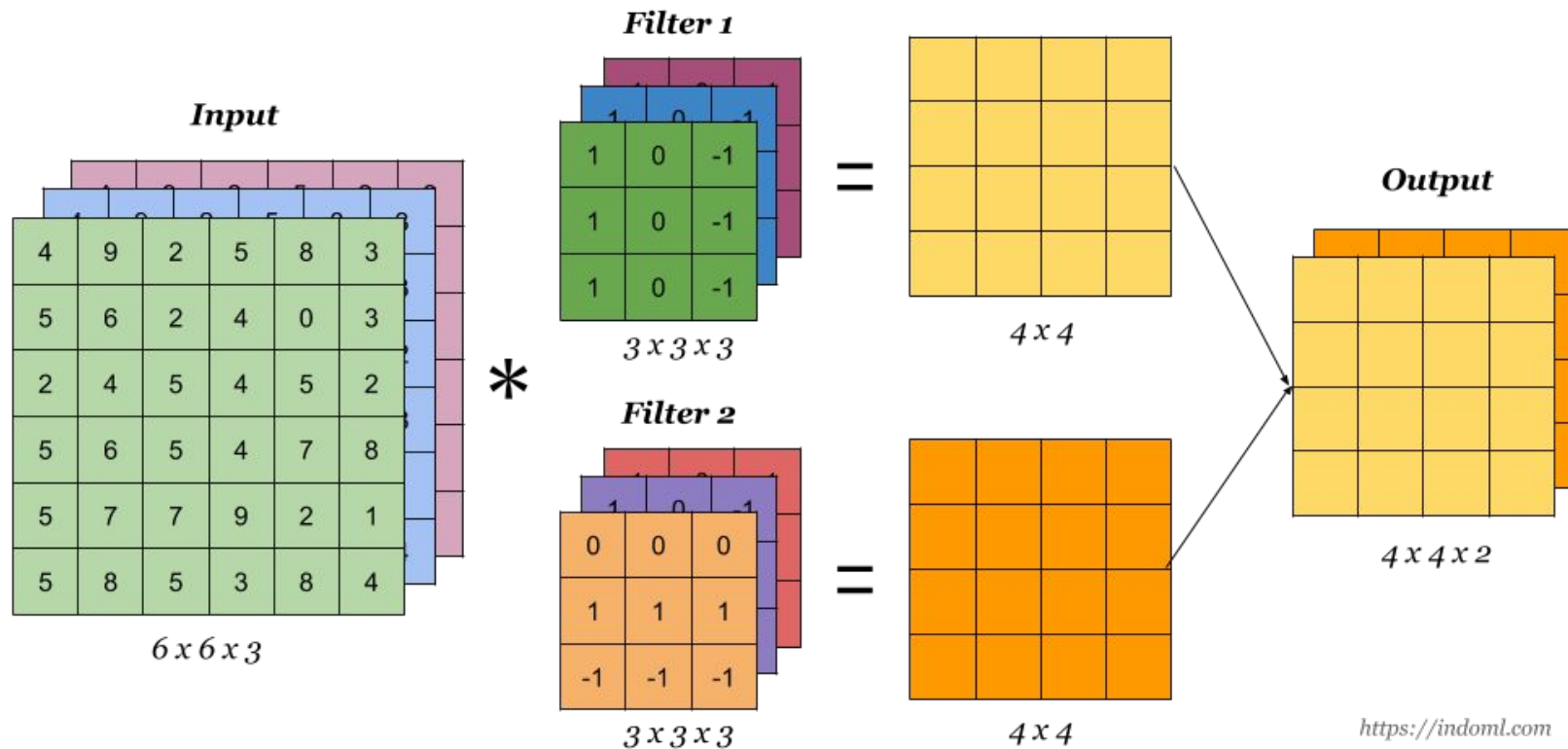
Mạng CNN phổ thông

Tóm tắt



Mạng CNN phổ thông

Convolutional layer



Mạng CNN phổ thông

Convolutional layer - Padding

0	0	0	0	0	0	0	0
0	3	3	4	4	7	0	0
0	9	7	6	5	8	2	0
0	6	5	5	6	9	2	0
0	7	1	3	2	7	8	0
0	0	3	7	1	8	3	0
0	4	0	4	3	2	2	0
0	0	0	0	0	0	0	0

$6 \times 6 \rightarrow 8 \times 8$

*

1	0	-1
1	0	-1
1	0	-1

3×3

=

-10	-13	1			
-9	3	0			

6×6

Sử dụng padding không làm giảm size bức ảnh

Mạng CNN phổ thông

Convolutional layer - Tính chất

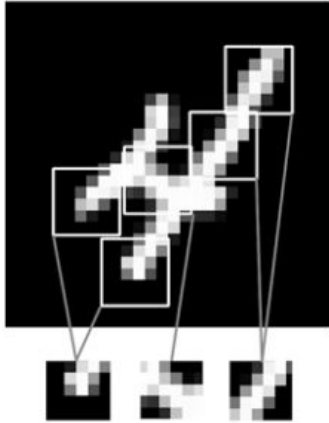


Figure 5.1 Images can be broken into local patterns such as edges, textures, and so on.

The patterns they learn are translation invariant.

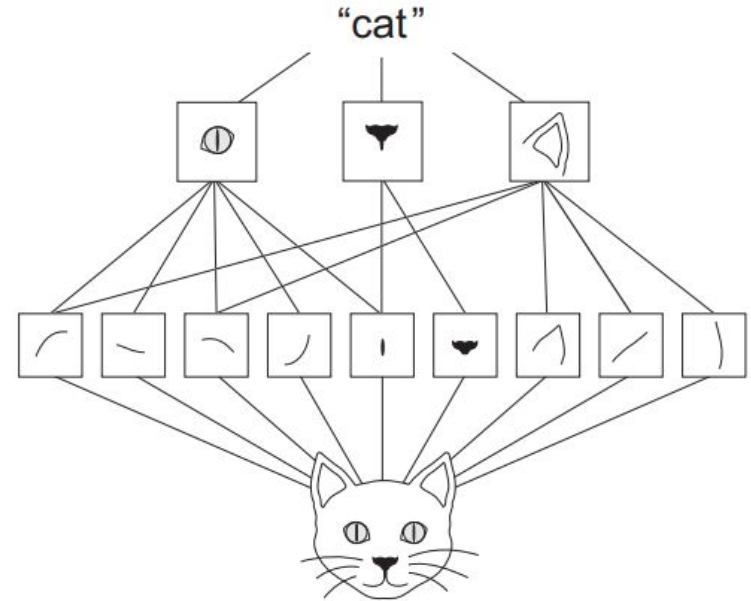


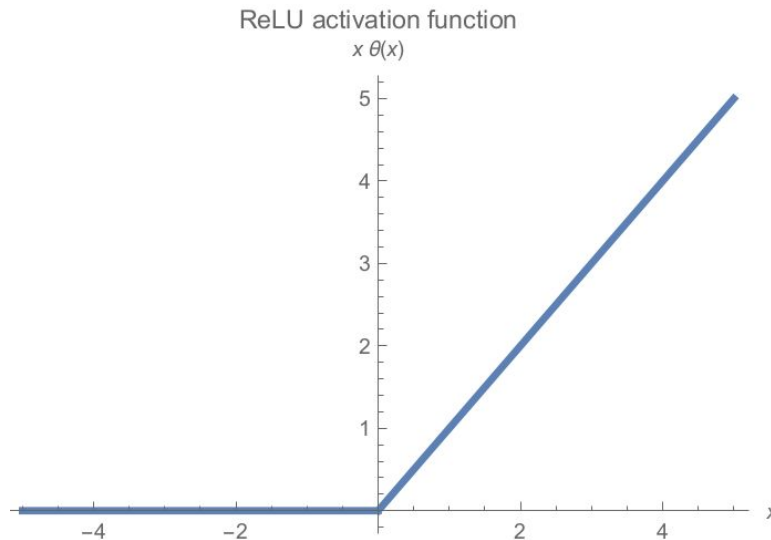
Figure 5.2 The visual world forms a spatial hierarchy of visual modules: hyperlocal edges combine into local objects such as eyes or ears, which combine into high-level concepts such as “cat.”

Mạng CNN phổ thông

Activation layer (ReLU)

- Hàm ReLU activation đơn giản để tính
- Đạo hàm là 1 với $x \geq 0$ nên không bị vanishing gradient.

(Vanishing gradient là hiện tượng đạo hàm của các layer đầu bị quá nhỏ dẫn đến việc khó khăn trong tối ưu hàm loss)



Mạng CNN phổ thông

Pooling layer

Tác dụng:

- Giảm số tham số
- Phù hợp với cấu trúc phân lớp (Layer sau sẽ nhìn vào một vùng rộng hơn)

Giải thích:

- Nếu chúng ta sử dụng 3x3 window thì đến layer thứ 3 chúng ta chỉ nhìn vào 7x7 pixel tương ứng với ảnh gốc
- Sử dụng pooling giống như việc tổng kết, giữ lại những đặc trưng tốt nhất
=> Tốt cho cấu trúc phân cấp

Max Pooling

29	15	28	184
0	100	70	38
12	12	7	2
12	12	45	6

2 x 2
pool size

100	184
12	45

Average Pooling

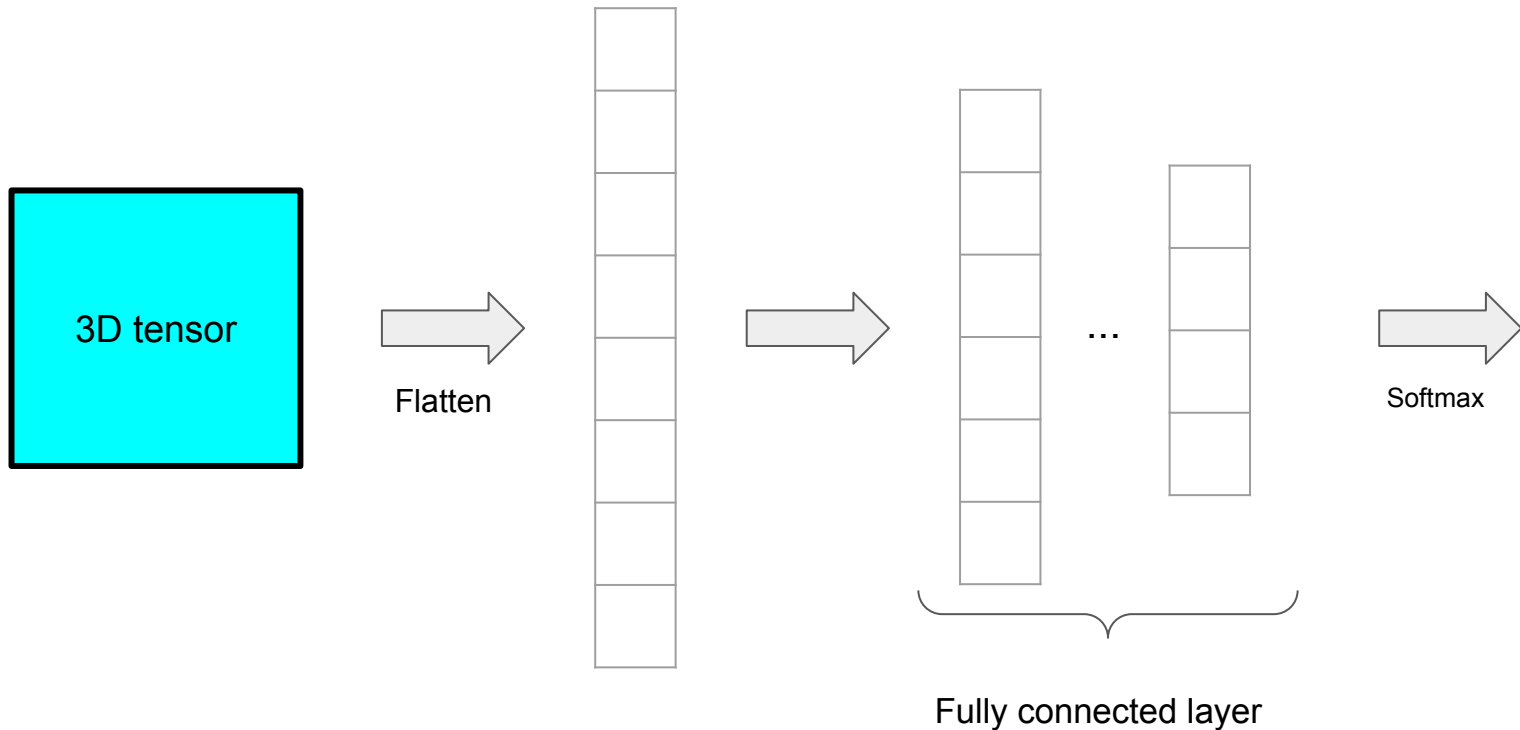
31	15	28	184
0	100	70	38
12	12	7	2
12	12	45	6

2 x 2
pool size

36	80
12	15

Mạng CNN phổ thông

Fully connected layer



Bài tập

- Suy nghĩ mạng CNN sẽ đạo hàm ngược như nào?
- Tìm hiểu thêm về các mạng CNN hiện đại.