



Bài 9: Convolutional Neural Network

Đội ngũ giảng dạy: - Hoàng Quý Phát
- Hoàng Minh Phương
- Huỳnh Chí Kiên
- Vũ Thái Duy
- Nguyễn Vinh Tiệp

Cố vấn khóa học: Huyền Nguyễn, Tiệp Vũ, Thăng Lương



Convolutional Neural Network

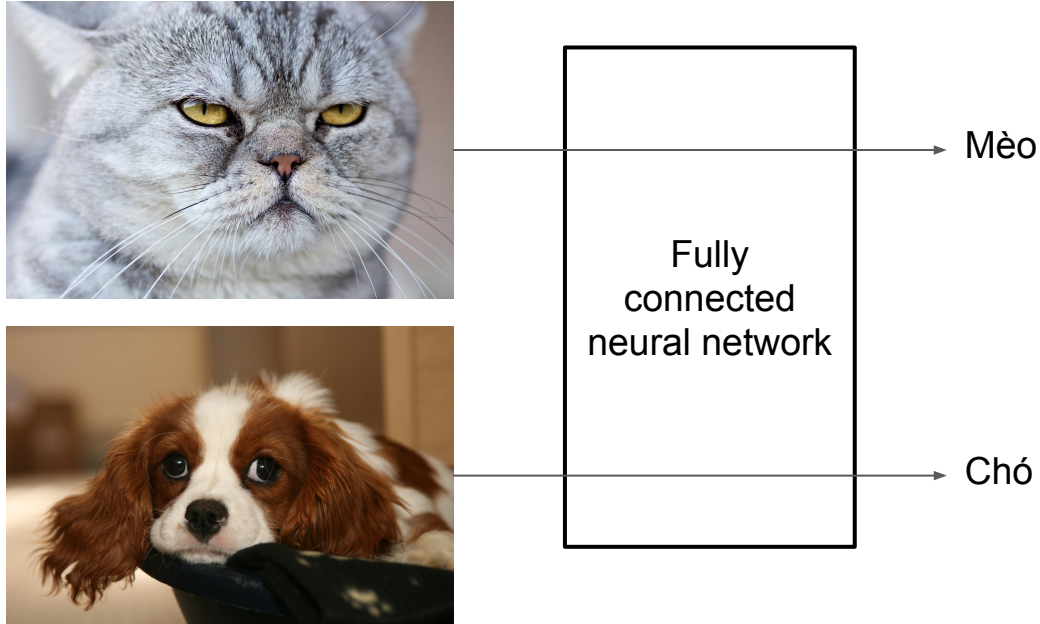
Outline

- Giới thiệu
- Phép tích chập 1D rời rạc
- Phép tích chập 2D rời rạc
- Mạng CNN
- Một số CNN phổ biến
- Các ứng dụng khác của CNN
- Hiển thị CNN để quan sát



Convolutional Neural Network

Giới thiệu





Convolutional Neural Network

Giới thiệu



=

1	123	32	4	221	19	6	2	8	12
255	19	6	2	8	32	1	6	7	99
99	4	5	1	123	32	4	221	12	7
32	1	6	7	4	33	6	253	54	12
19	6	2	8	19	6	2	8	123	44
11	14	44	33	25	32	1	6	7	44
12	8	19	6	2	8	78	99	3	33
123	32	4	221	66	24	32	1	6	7

flatten

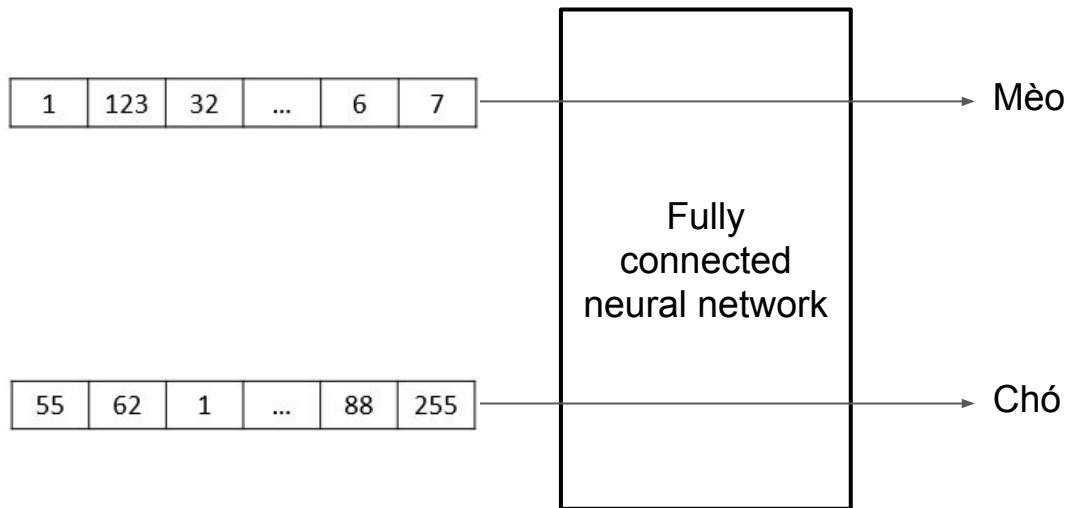
1	123	32	...	6	7
---	-----	----	-----	---	---

1 x (H x W)

Ma trận với kích thước
H x W

Convolutional Neural Network

Giới thiệu

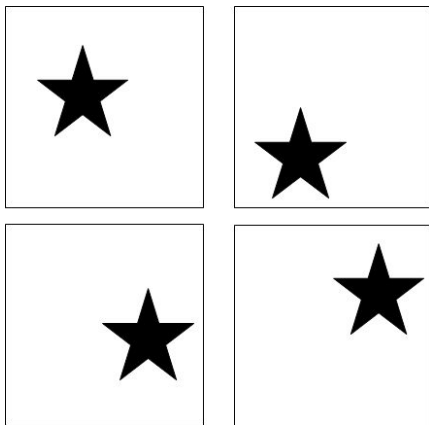




Convolutional Neural Network

Giới thiệu

- Các vấn đề phát sinh khi dùng mạng kết nối đầy đủ để phân loại hình ảnh:
 - Dễ bị overfit.
 - Lượng tham số cần chứa quá lớn. Giả sử ảnh đầu vào kích thước $256 \times 256 \times 3$, lớp ẩn đầu tiên có 256 nút ẩn \rightarrow Lớp ẩn này sẽ cần tất cả 50,331,648 tham số.
 - Đối tượng chính trong ảnh chỉ cần thay đổi ít về vị trí là mạng có thể phân loại sai.

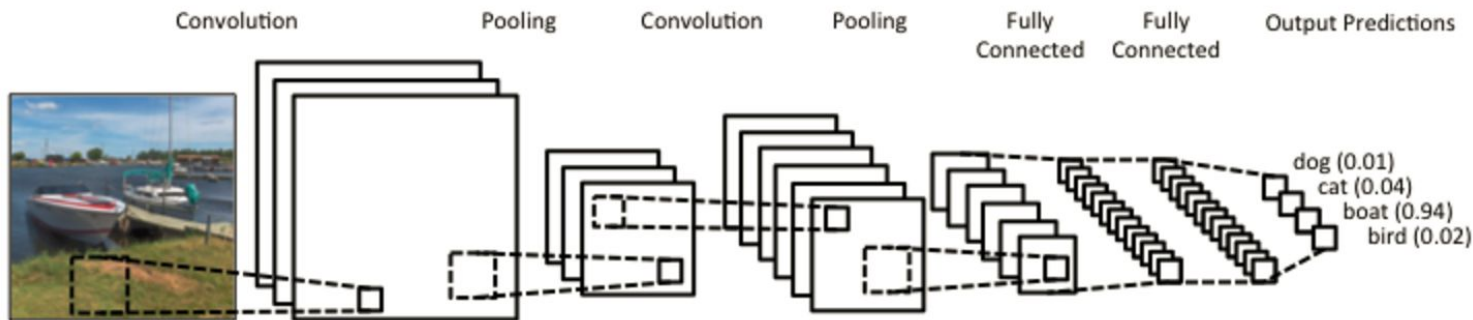




Convolutional Neural Network

Giới thiệu

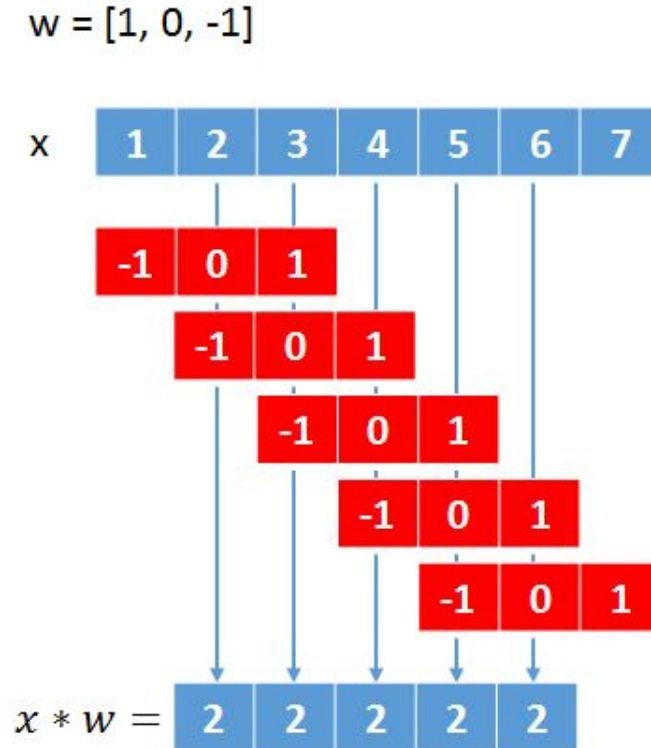
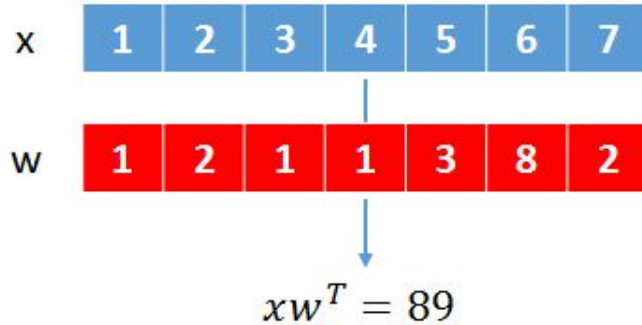
- Convolutional neural network là mạng được thiết kế với các đặc thù:
 - Lượng tham số ít hơn mạng kết nối đầy đủ nhiều lần.
 - Có tính kháng phép dịch chuyển (translational invariance).





Convolutional Neural Network

Phép tích chập 1D rời rạc





Convolutional Neural Network

Phép tích chập 1D rời rạc

- Định nghĩa:

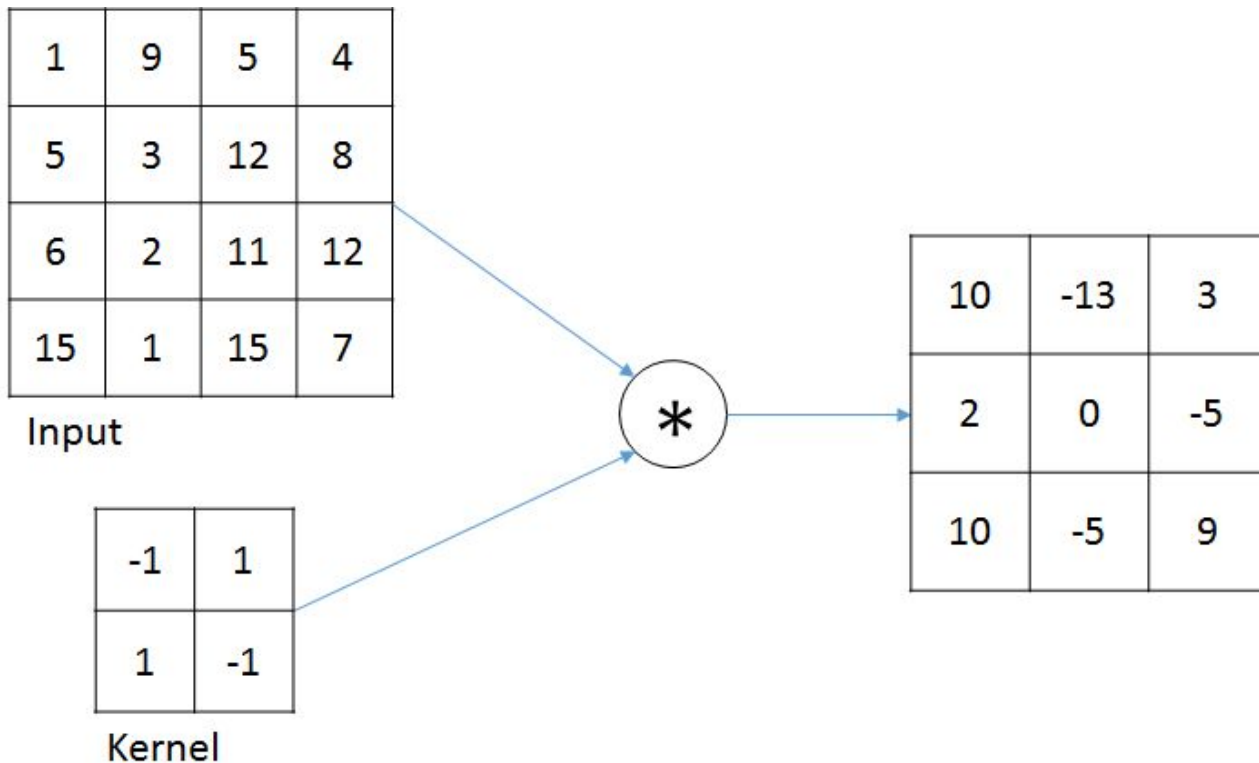
$$(f * g)[n] = \sum_{m=-\infty}^{+\infty} f[m]g[n - m]$$

- Đặc tính: gọi $g1$ là hàm đối xứng của g qua gốc tọa độ, tích chập giữa hàm f và g phản ánh mối tương quan giữa f và hàm $g1$.
- Một số từ khóa tương đồng khi nhắc đến g :
 - Cửa sổ (window)
 - Bộ lọc (filter)
 - Kernel
 - Mask
 - Filter mask



Convolutional Neural Network

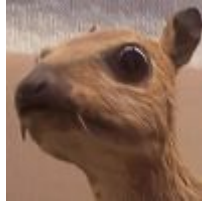
Phép tích chập 2D rời rạc





Convolutional Neural Network

Phép tích chập 2D rời rạc



$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$



Identity

Sobel X

Sobel Y

Laplacian

Sharpen

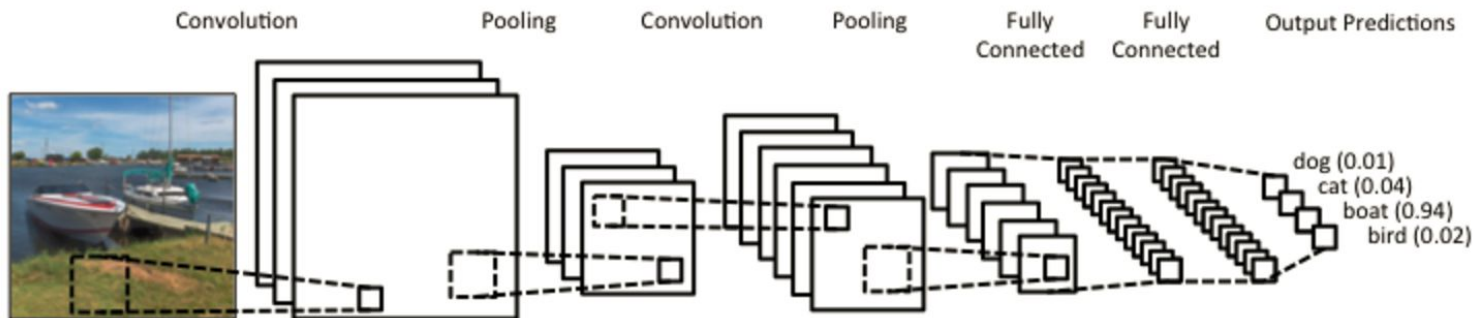
- Lưu ý: Kích thước của kernel không nhất thiết phải là 3 x 3, mà có thể là 5 x 5, 11 x 11, v.v.

Convolutional Neural Network

Mạng CNN



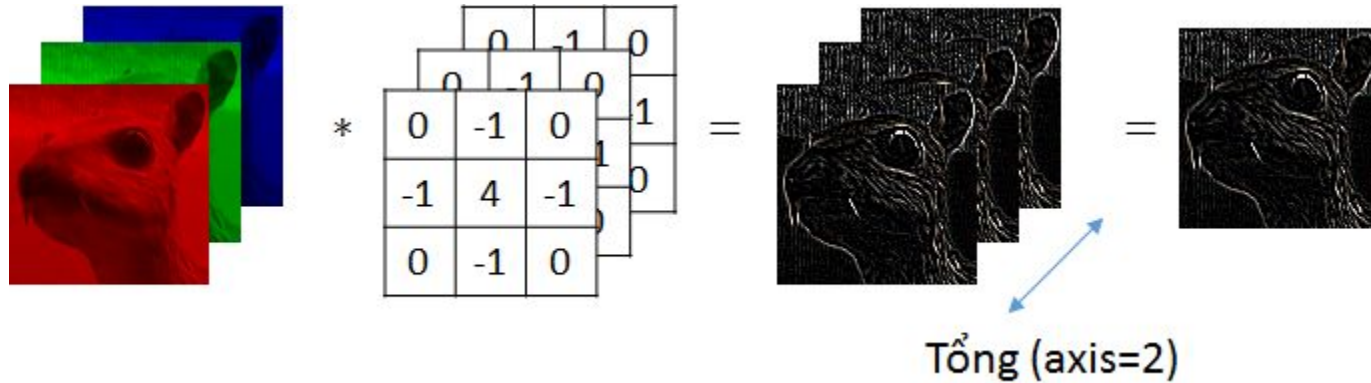
- Một CNN thường gồm có các thành phần:
 - Các lớp convolution (convolutional layer)
 - Các hàm kích hoạt (activation function)
 - Các lớp pooling
 - Các lớp fully-connected





Convolutional Neural Network

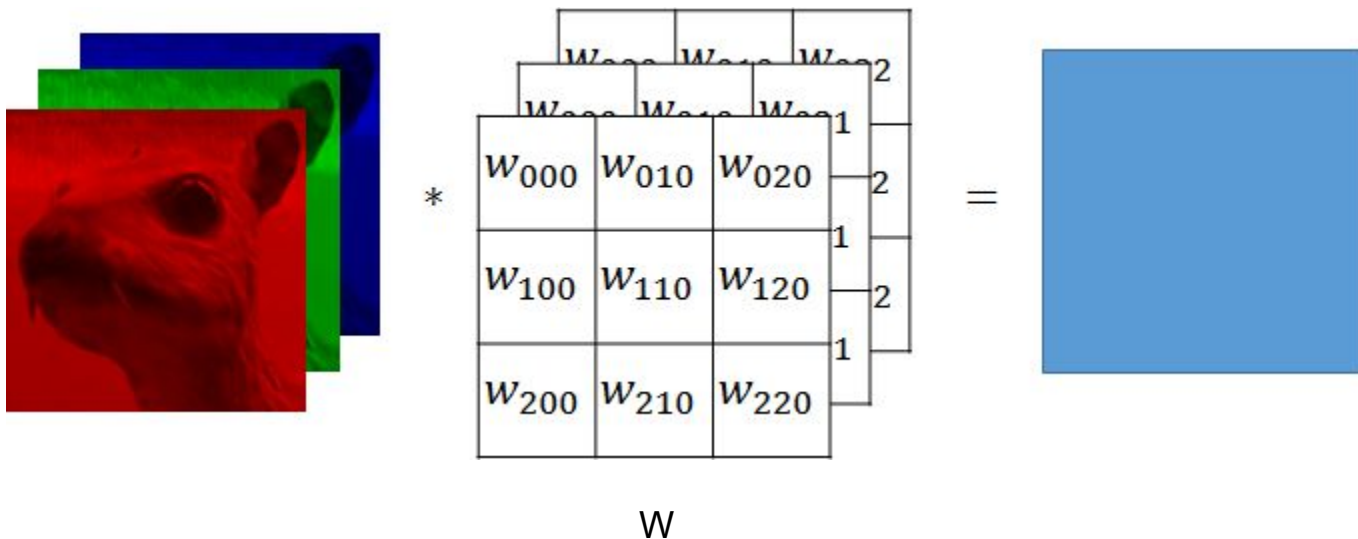
Mạng CNN - Convolutional Layer





Convolutional Neural Network

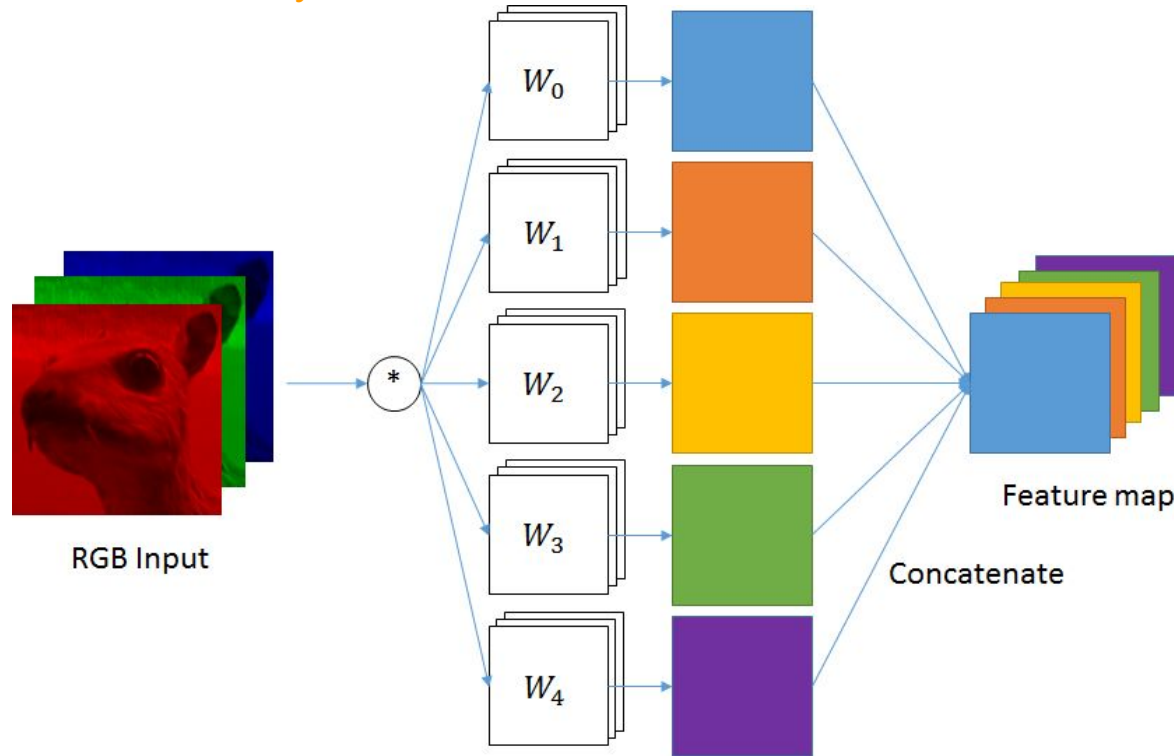
Mạng CNN - Convolutional Layer





Convolutional Neural Network

Mạng CNN - Convolutional Layer

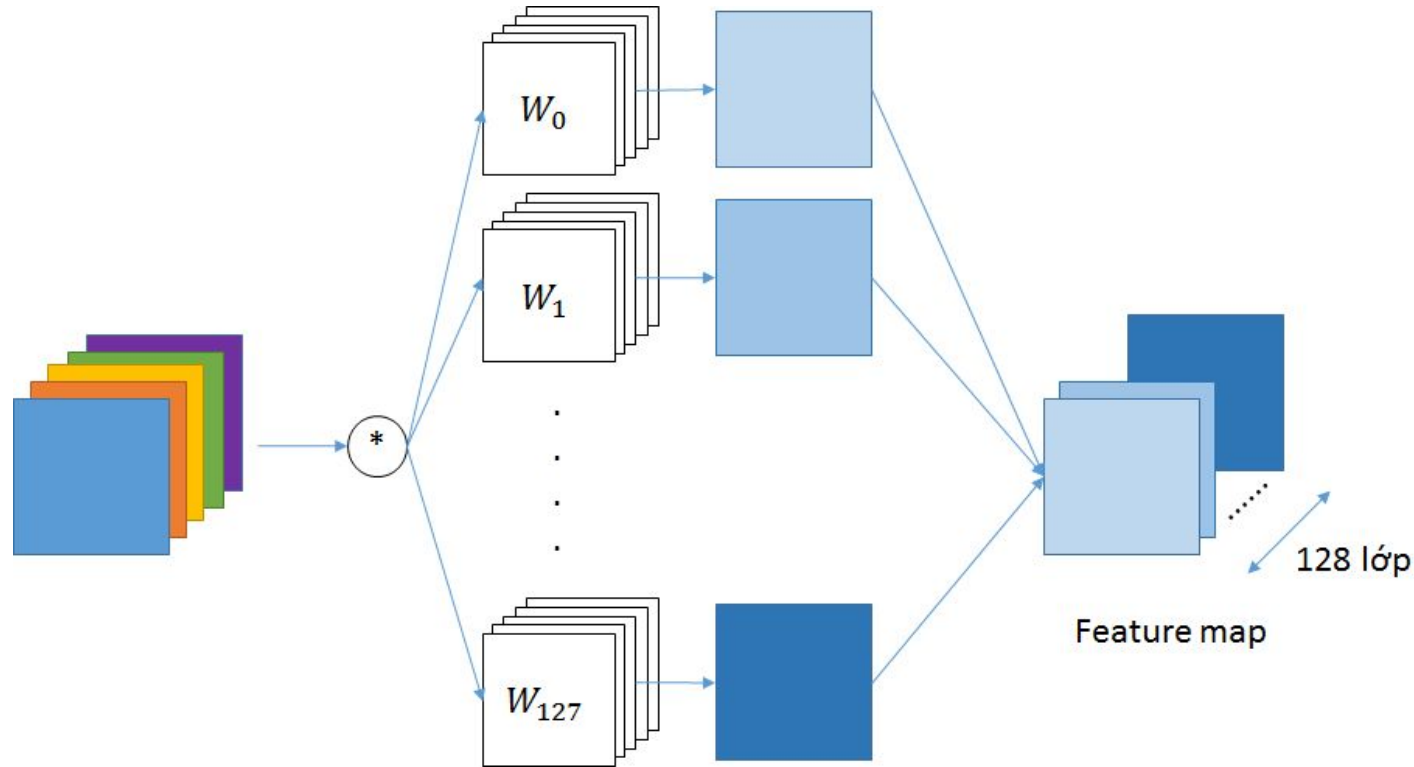


Tensor W có kích thước là $3 \times 3 \times 3 \times 5$ ($H \times W \times N_{in} \times N_{out}$)



Convolutional Neural Network

Mạng CNN - Convolutional Layer

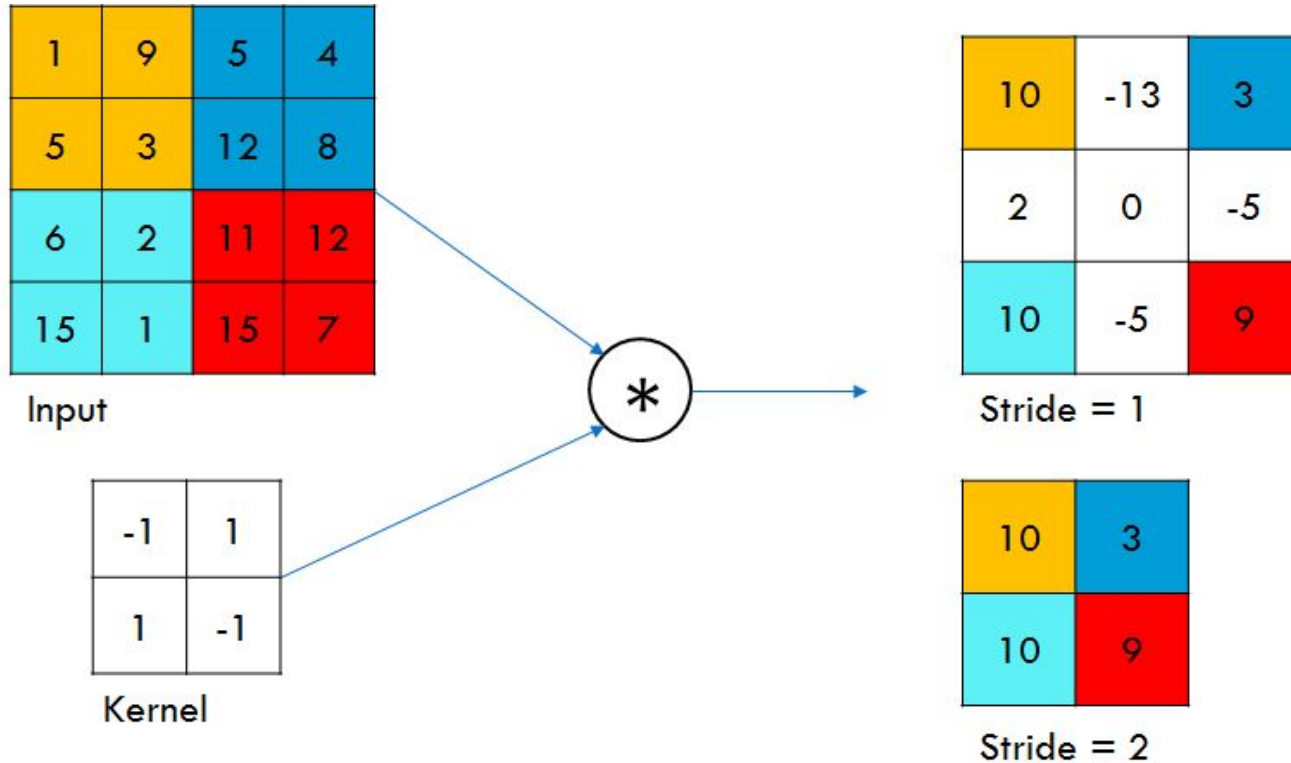


Tensor W có kích thước là $3 \times 3 \times 5 \times 128$ ($H \times W \times N_{in} \times N_{out}$)



Convolutional Neural Network

Mạng CNN - Stride





Convolutional Neural Network

Mạng CNN - Padding

0	0	0	0	0
0	1	9	5	0
0	5	3	12	0
0	6	2	11	0
0	0	0	0	0

Zero padding kích thước
(1,1)

0	0	0	0	0
0	0	0	0	0
0	1	9	5	0
0	5	3	12	0
0	6	2	11	0
0	0	0	0	0
0	0	0	0	0

Zero padding kích thước
(2,1)

1	1	1	9	5	5	5
1	1	1	9	5	5	5
1	1	1	9	5	5	5
5	5	5	3	12	12	12
6	6	6	2	11	11	11
6	6	6	2	11	11	11
6	6	6	2	11	11	11

Replicate padding kích thước
(2,2)

3	5	5	3	12	12	3
9	1	1	9	5	5	9
9	1	1	9	5	5	9
3	5	5	3	12	12	3
2	6	6	2	11	11	2
2	6	6	2	11	11	2
3	5	5	3	12	12	3

Symmetric padding kích thước
(2,2)



Convolutional Neural Network

Mạng CNN - Activation function

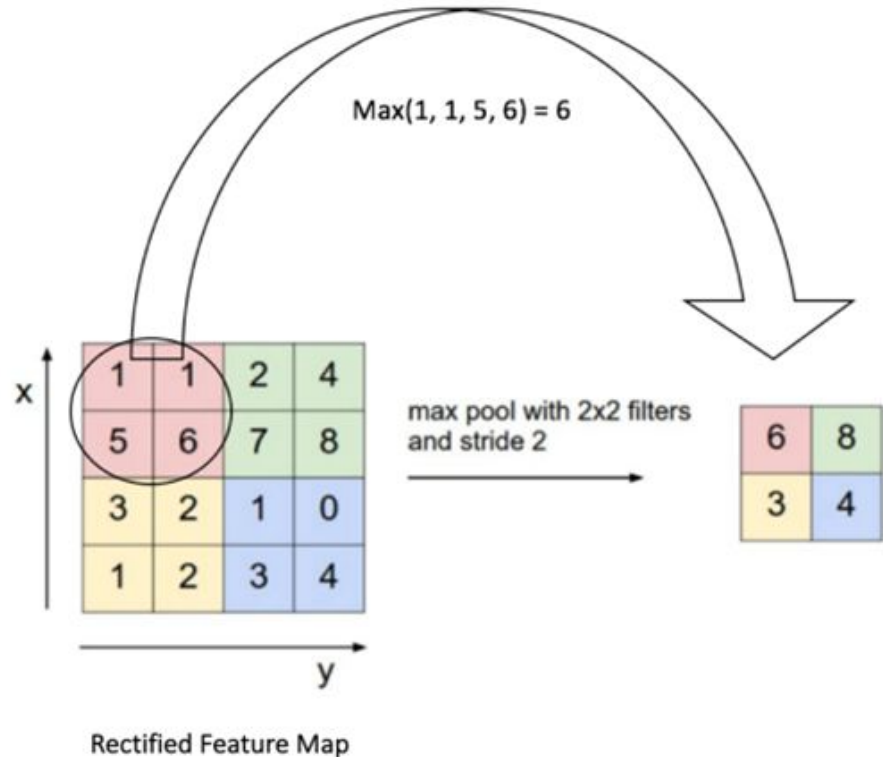
- Như một NN bình thường, ta có thể hoàn toàn dùng các hàm sau làm hàm kích hoạt sau mỗi lớp convolution:
 - ReLU
 - Tanh
 - Sigmoid



Convolutional Neural Network

Mạng CNN - Pooling

- Có hai loại pooling chính trong CNN: max-pooling và average-pooling.
- Có thể hiểu max-pooling là một dạng kernel mà output là giá trị lớn nhất của các phần tử input nằm trong kernel đó.
- Thông thường, số stride của tầng pooling sẽ chính bằng kích thước kernel.

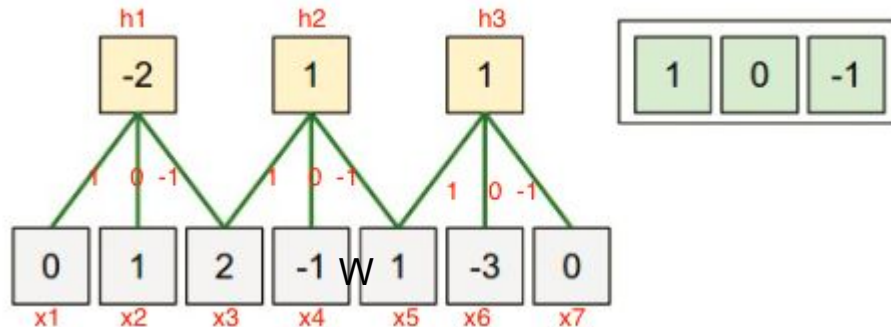




Convolutional Neural Network

Mạng CNN

- Các tầng convolution có hai đặc điểm chính khác với mạng fully-connected:
 - Weight sharing: các nút output khác nhau dùng chung một bộ tham số (w).
 - Sparsely-connected (kết nối thưa): một nút đầu ra chỉ liên kết với 1 số nút đầu vào chứ không phải toàn bộ.



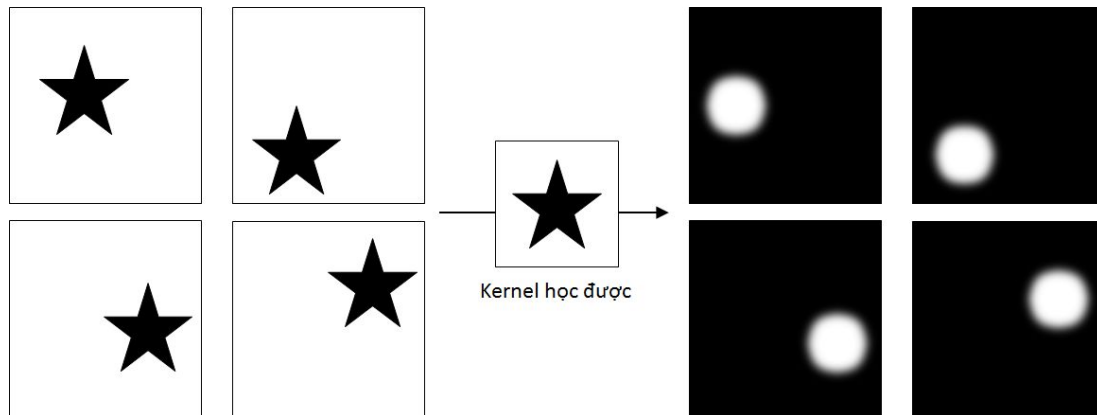
Weight sharing: kernel $[1, 0, -1]$ được dùng để tính cho cả ba nút $h1, h2$ và $h3$.



Convolutional Neural Network

Mạng CNN - Thể mạnh

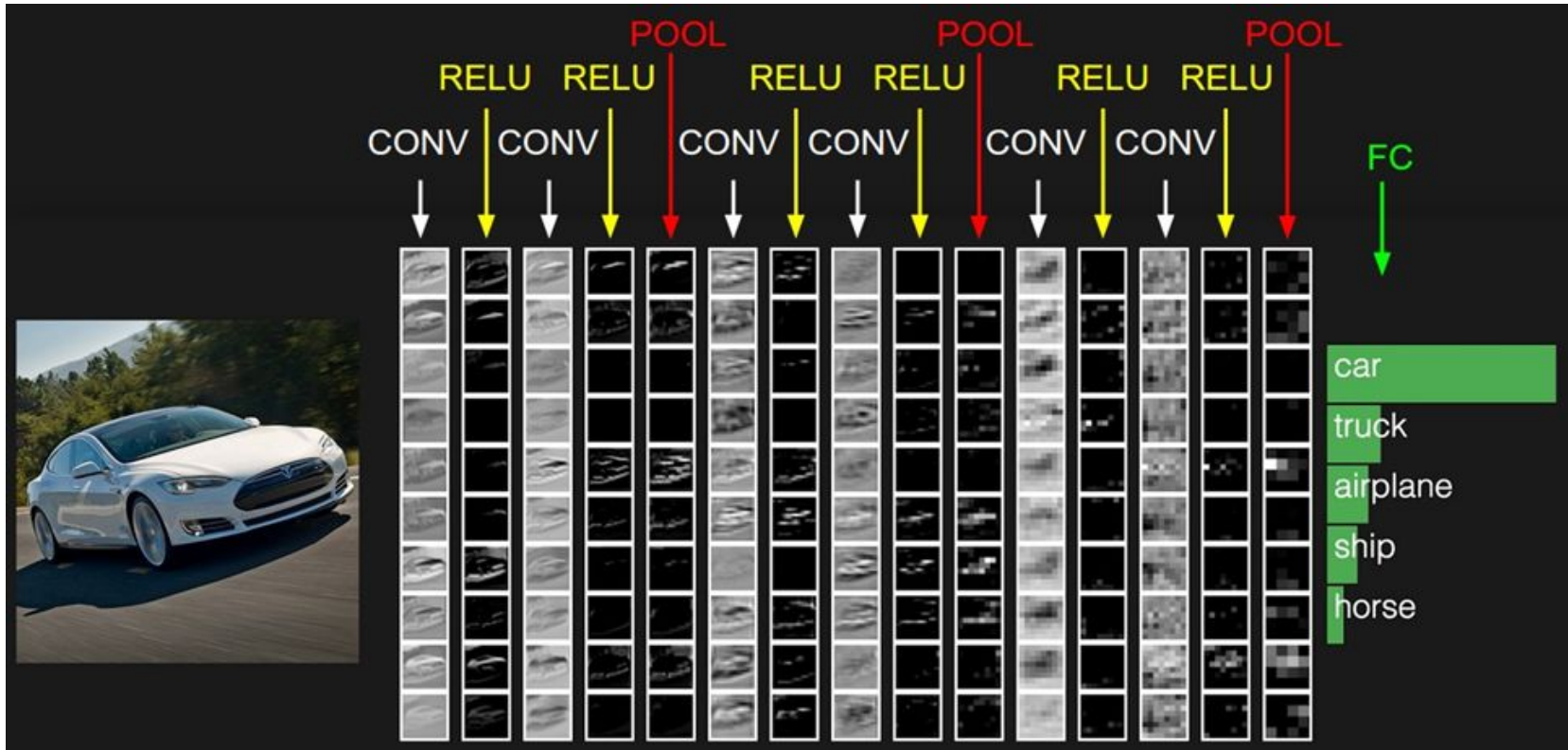
- CNN hoạt động tốt hơn fully-connected NN vì:
 - Phép tích chập có tính translational invariance.
 - Lượng tham số ít hơn rất nhiều.
 - Vì lượng tham số ít hơn nên khó overfit hơn.
- CNN thường hoạt động tốt trên các loại data mà giữa các đặc trưng có quan hệ về mặt không gian (ảnh, sóng âm, v.v.)





Convolutional Neural Network

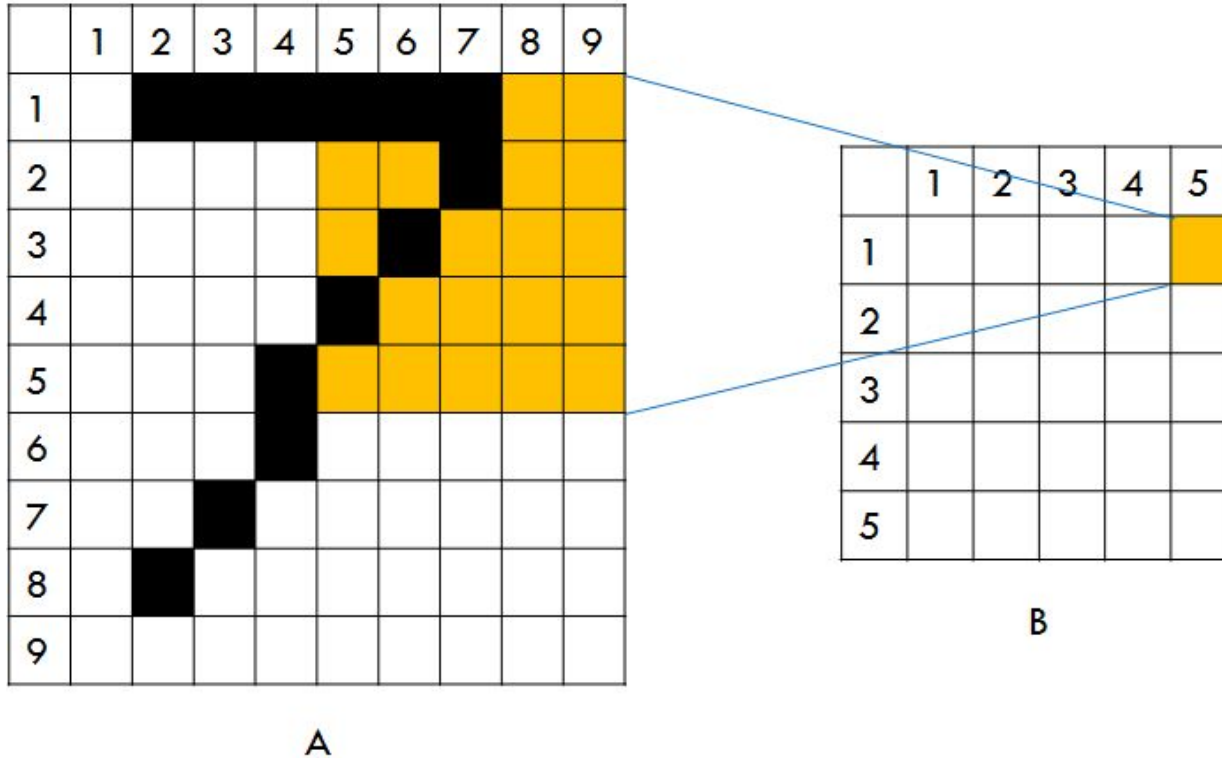
Mạng CNN





Convolutional Neural Network

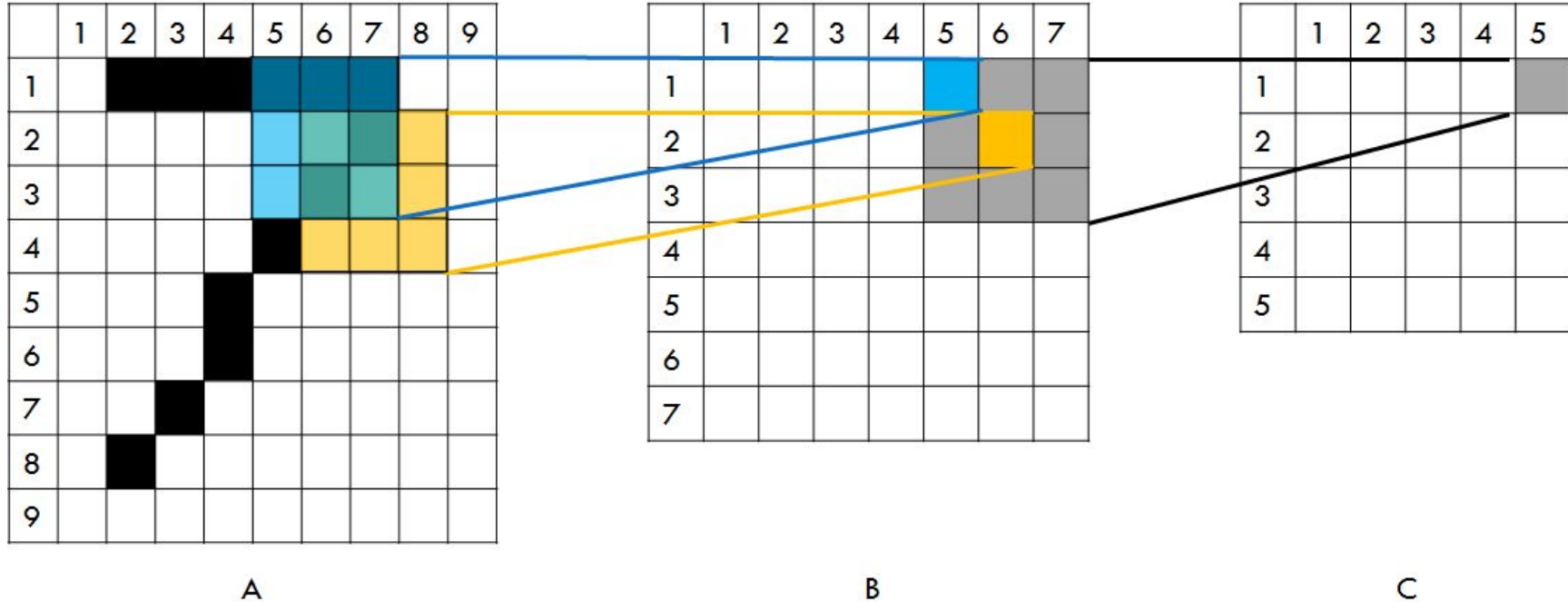
Mạng CNN - Receptive field





Convolutional Neural Network

Mạng CNN - Effective receptive field



- Khi thiết kế một CNN mới (nếu thực sự cần!), cần lưu ý sao cho effective receptive field bao phủ được các chi tiết cần thiết trong ảnh đầu vào



Convolutional Neural Network

Hiển thị CNN để quan sát

