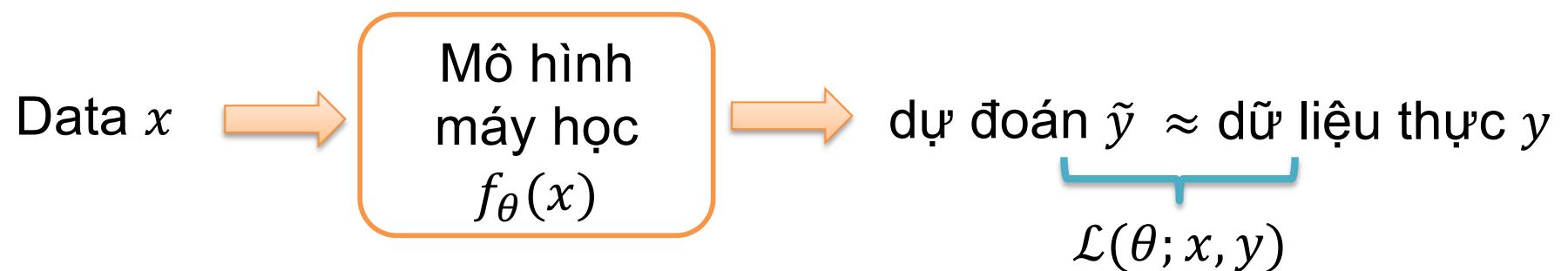


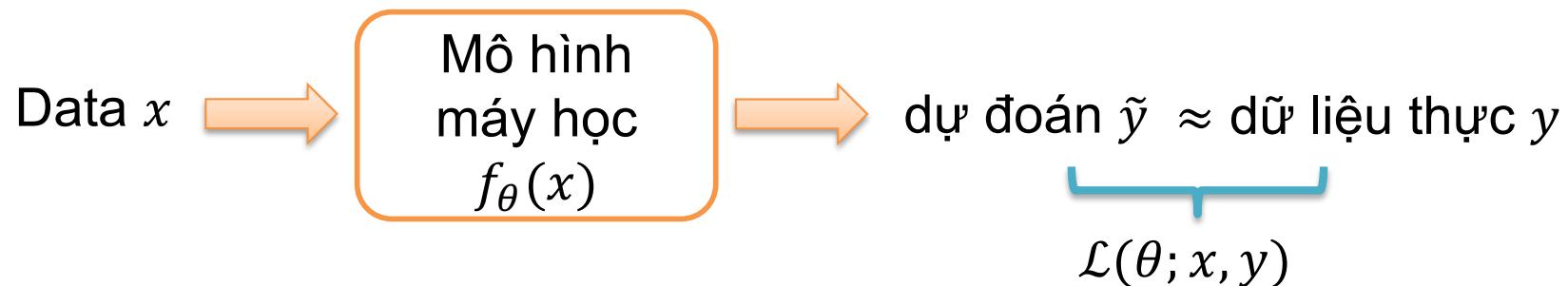
CÁC KỸ THUẬT HỌC
SÂU VÀ ỨNG DỤNG

MẠNG CONVOLUTIONAL
NEURAL NETWORK

Mô hình học tổng quát



Mô hình học tổng quát



- Huấn luyện mô hình từ tập dữ liệu huấn luyện $\{(x_i, y_i)\}_{i=1..n}$
- Hay, tìm tham số θ của mô hình f để $\tilde{y} \approx y$
- Hay, tìm tham số θ để hàm độ lỗi $\mathcal{L}(\theta; x, y)$ nhỏ nhất

Nội dung

- Giới thiệu
- Convolutional neural networks (CNN)
- Cài đặt CNN với Tensorflow

Giới thiệu - Các loại ảnh phổ biến

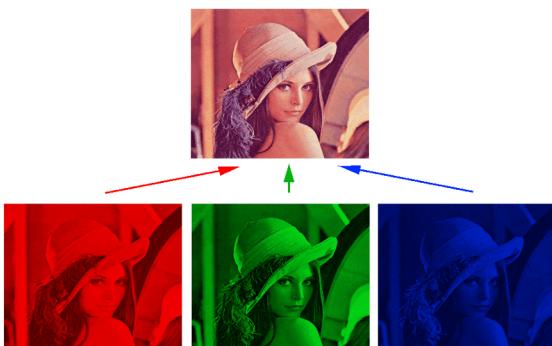
Ảnh xám



0	120	3	0	6
100	0	43	2	128
2	1	255	0	1
34	0	4	5	7
1	1	56	0	1

height

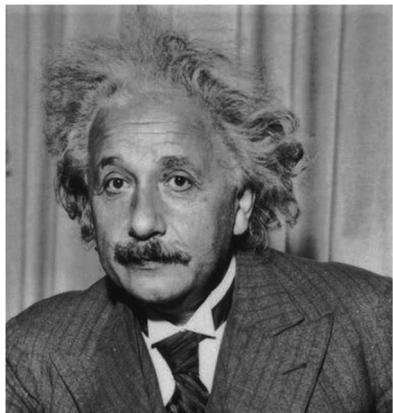
Ảnh RGB



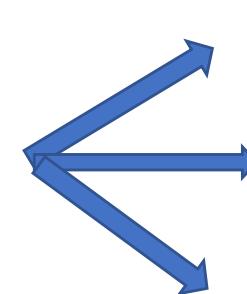
depth

0	120	3	0	6			
100	0	120	3	0	6		
2	100	0	120	3	0	6	
34	2	100	0	43	2	128	
1	34	2	1	255	0	1	
	1	34	0	4	5	7	
		1	1	56	0	1	

Giới thiệu - Bài toán phân loại cho ảnh



Neural
networks

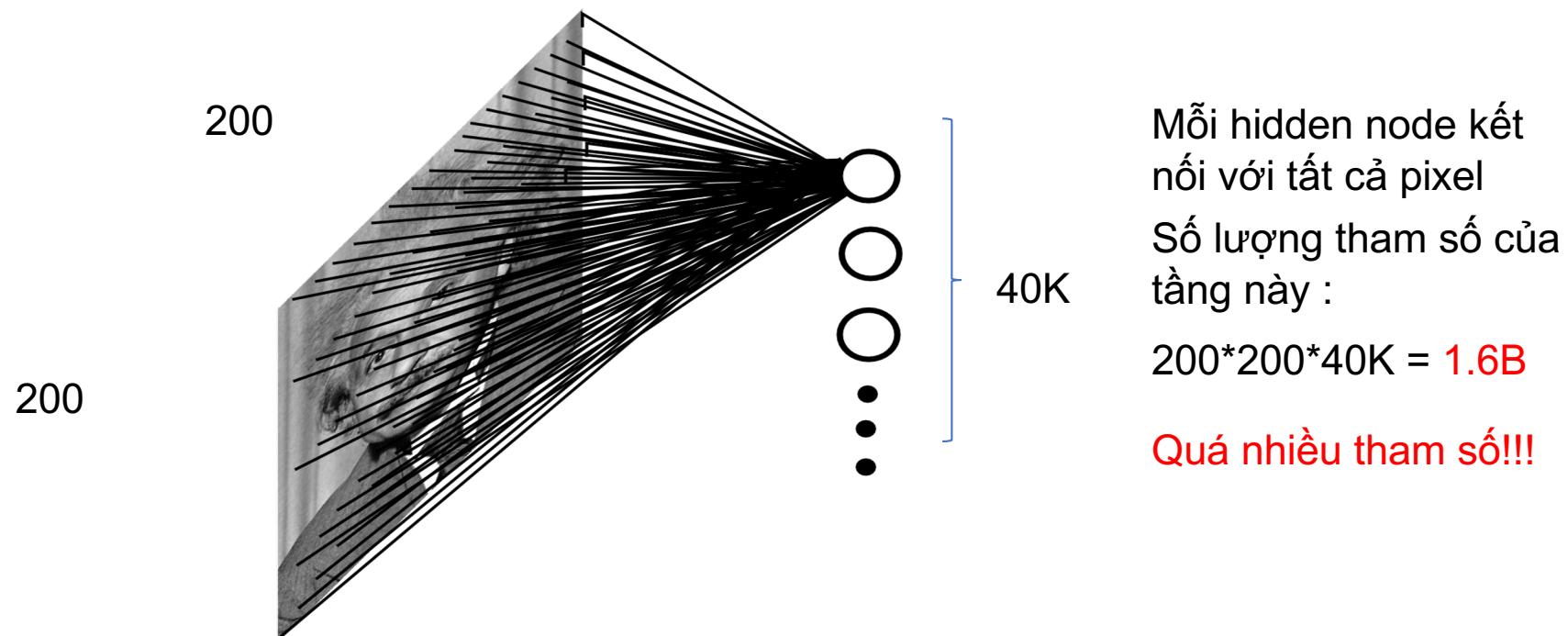


Xe cộ

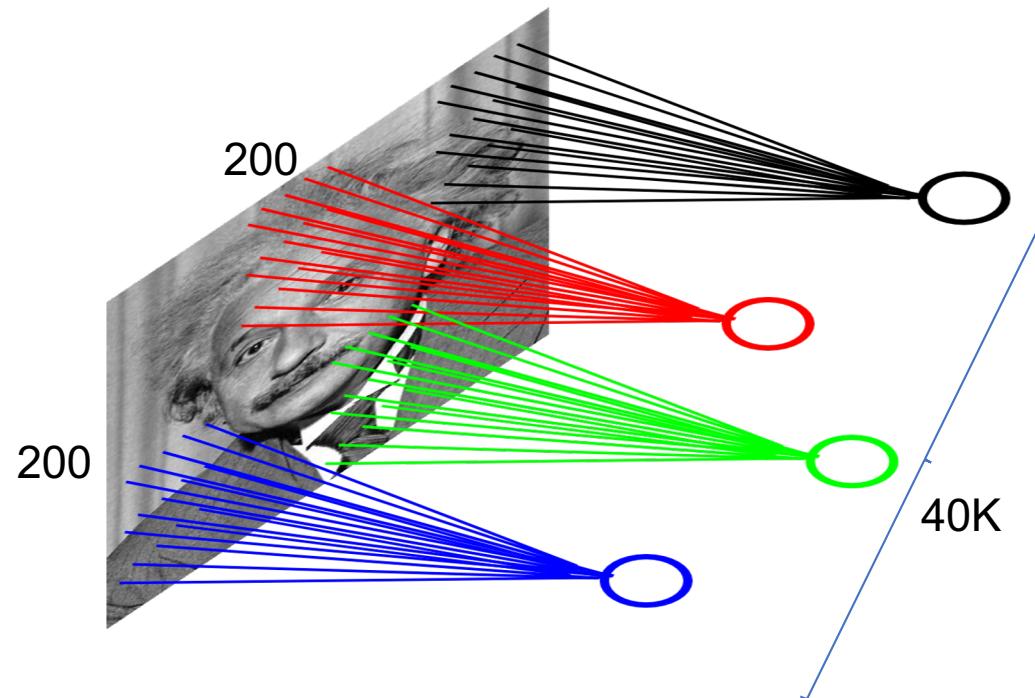
Nhà cửa

Con người

Giới thiệu - Mạng neuron fully connected



Giới thiệu - Mạng neuron locally connected

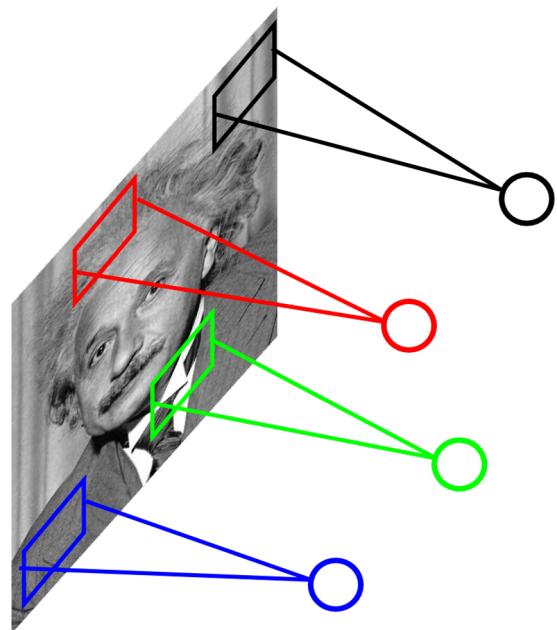


Mỗi hidden node chỉ
kết nối với một vùng
pixel **cố định** có kích
thước 10×10

Số lượng tham số của
tầng :
 $200 \times 200 \times 10 \times 10 = 4M$

Số lượng tham số giảm đáng
kể so với fully-connected!

Giới thiệu - Mạng neuron locally connected

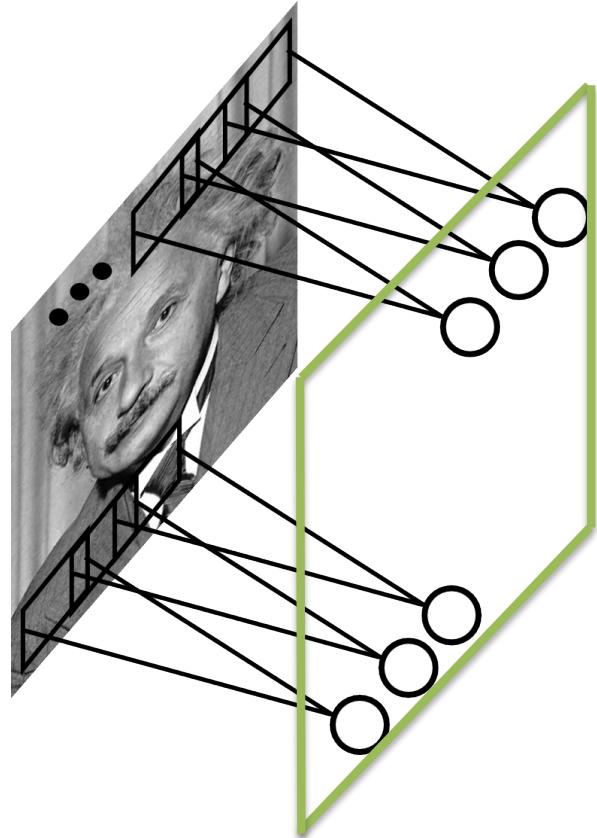


Hạn chế: Chỉ phù hợp với ảnh **có bối cảnh rõ ràng**, chẳng hạn như ảnh khuôn mặt người

Làm thế nào để áp dụng cho những ảnh thông thường?

- Mỗi hidden node cần kết nối nhiều vùng pixel trên một bức ảnh
- **Chia sẻ tham số**

Mạng neuron weight-sharing locally connected



Tham số **được chia sẻ** cho các vùng ảnh khác nhau

Phép convolution

Phép Convolution

- Phép biến đổi Convolution:
 - Là một phép biến đổi tuyến tính
 - Rút trích đặc trưng hình ảnh (feature map)



*

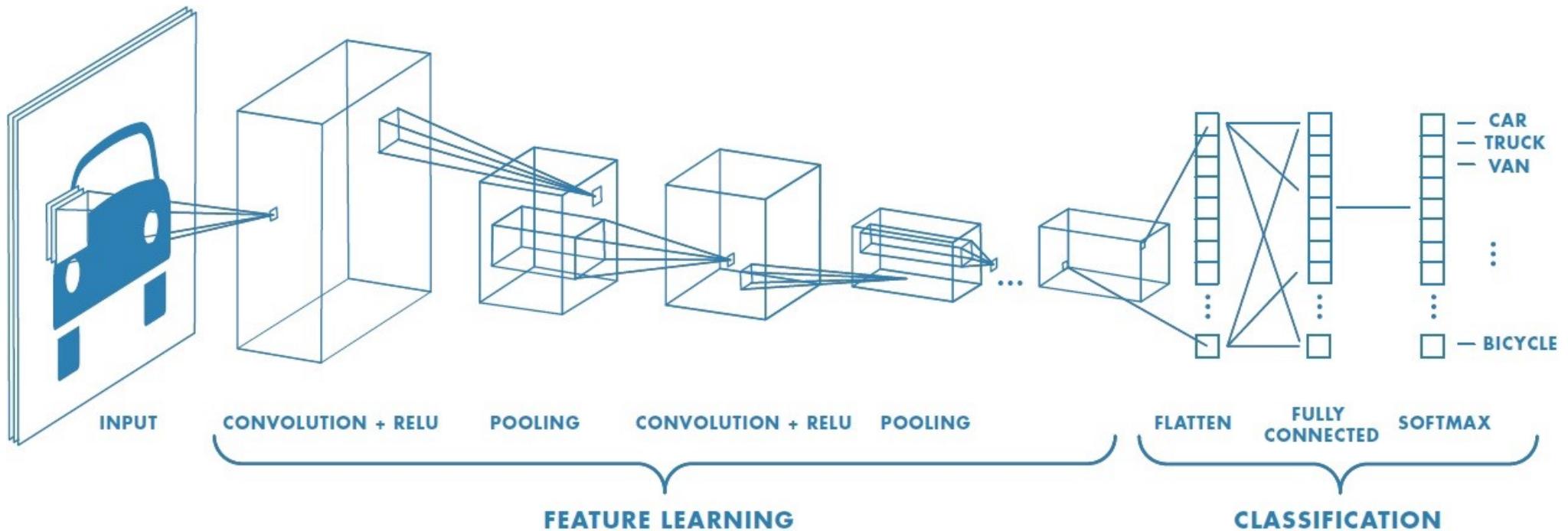
1	0	-1
2	0	-2
1	0	-1



Nội dung

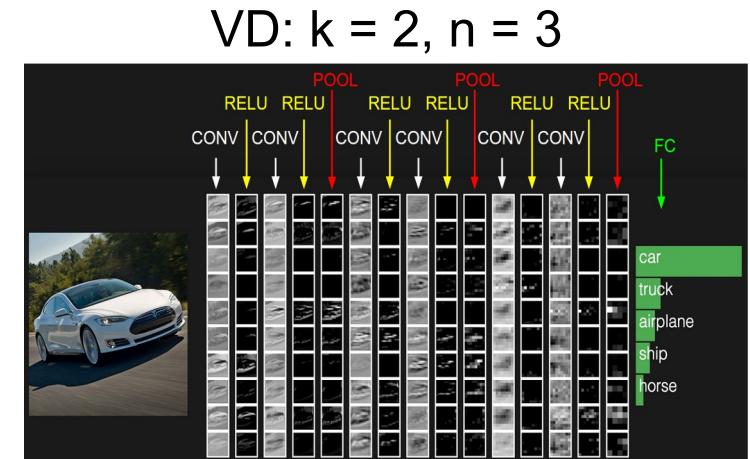
- Giới thiệu
- Convolutional neural networks (CNN)
- Cài đặt CNN với Tensorflow

CNN - Kiến trúc phổ dụng



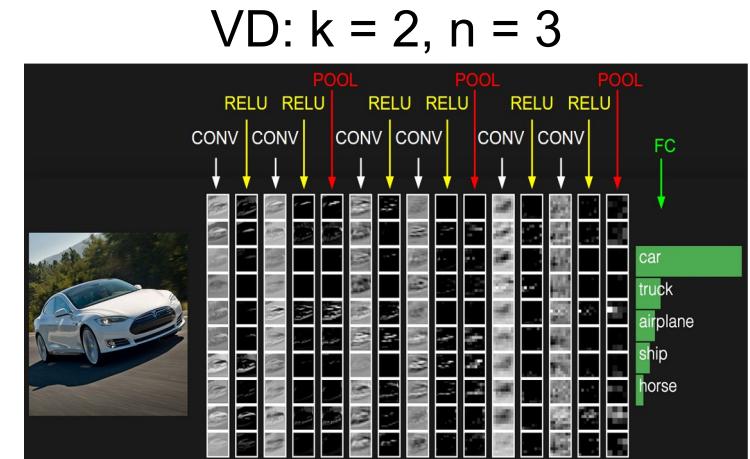
CNN - Kiến trúc phổ dụng

- Tầng convolution
 - Tầng activation (ReLU)
 - Tầng pooling (optional)
- } Lặp k lần
- Tầng fully connected
- } Lặp n lần



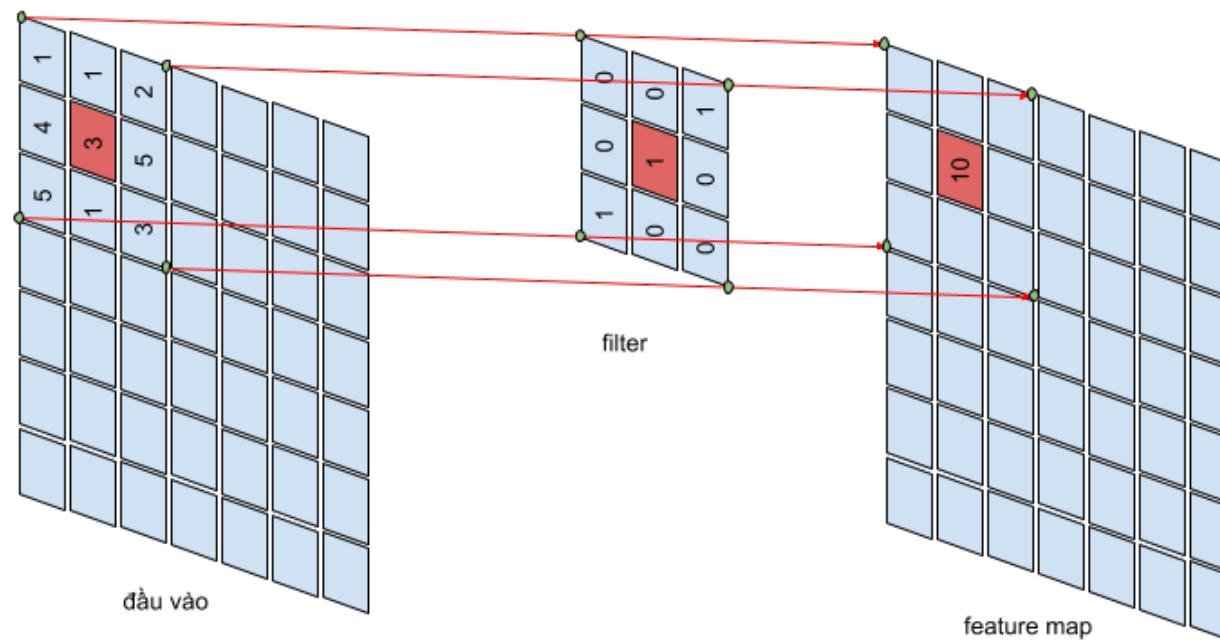
CNN - Kiến trúc phổ dụng

- Tầng convolution
 - Tầng activation (ReLU)
 - Tầng pooling (optional)
- } Lặp k lần
- Tầng fully connected
- } Lặp n lần



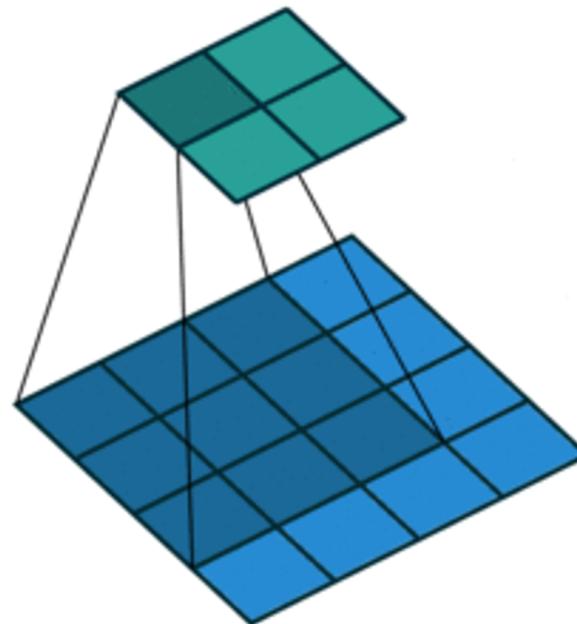
Tầng Convolution - 2D convolution

- Phép convolution tại một vị trí của bức ảnh



Tầng Convolution - 2D convolution

- Trượt filter **theo hai chiều** rộng và cao của tấm hình



Tầng Convolution - 2D convolution

- Trượt filter **theo hai chiều** rộng và cao của tấm hình

Ảnh 5x5

2	0	3	0	1
1	0	4	2	1
3	1	1	0	1
1	0	4	5	0
0	1	2	0	1

*

Filter 3x3

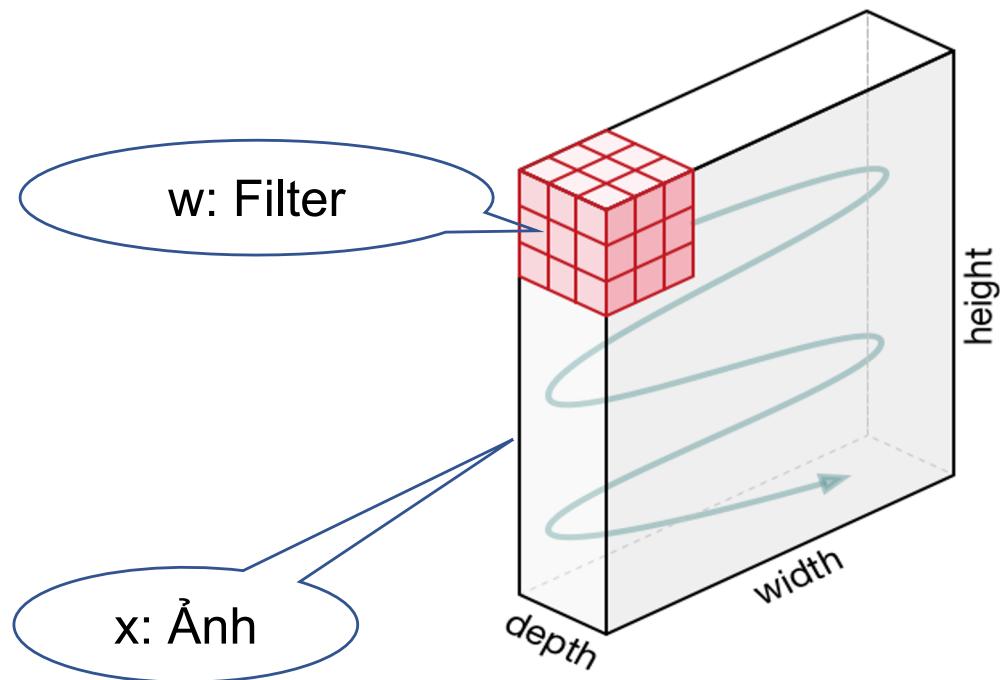
0	1	0
0	1	0
0	1	0

=

Feature map 3x3

1	8	2
1	9	7
2	7	5

Tầng Convolution - 2D Convolution cho ảnh RGB



Tầng Convolution - Stride

- Stride (S) là **độ dài bước trượt** của filter

Ảnh 5x5				
2	0	3	0	1
1	0	4	2	1
3	1	1	0	1
1	0	4	5	0
0	1	2	0	1

Filter 3x3

0	1	0
0	1	0
0	1	0

*

Feature map 3x3

1	8	2
1	9	7
2	7	5

S = 1

Feature map 2x2

1	2
2	5

S = 2

Tầng Convolution - Padding

- Khi biến đổi convolution:
 - Thông tin ở biên bức ảnh bị mất, và
 - Kích thước của ảnh giảm nhanh

→ Padding

Ảnh 3x3

0	4	2
1	1	0
3	4	5

*

Filter 3x3

0	1	0
0	1	0
0	1	0

=

Feature map 1x1

9

Tầng Convolution - Padding

Ảnh 3x3 được thêm pad

0	0	0	0	0
0	0	4	2	0
0	1	1	0	0
0	3	4	5	0
0	0	0	0	0

Filter 3x3

0	1	0
0	1	0
0	1	0

*

Feature map 3x3

1	5	2
4	9	7
4	5	6

=

Tầng Convolution - Padding

- Một số loại padding khác

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

Zero padding kích thước
(1,1)

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

Zero padding kích thước
(2,1)

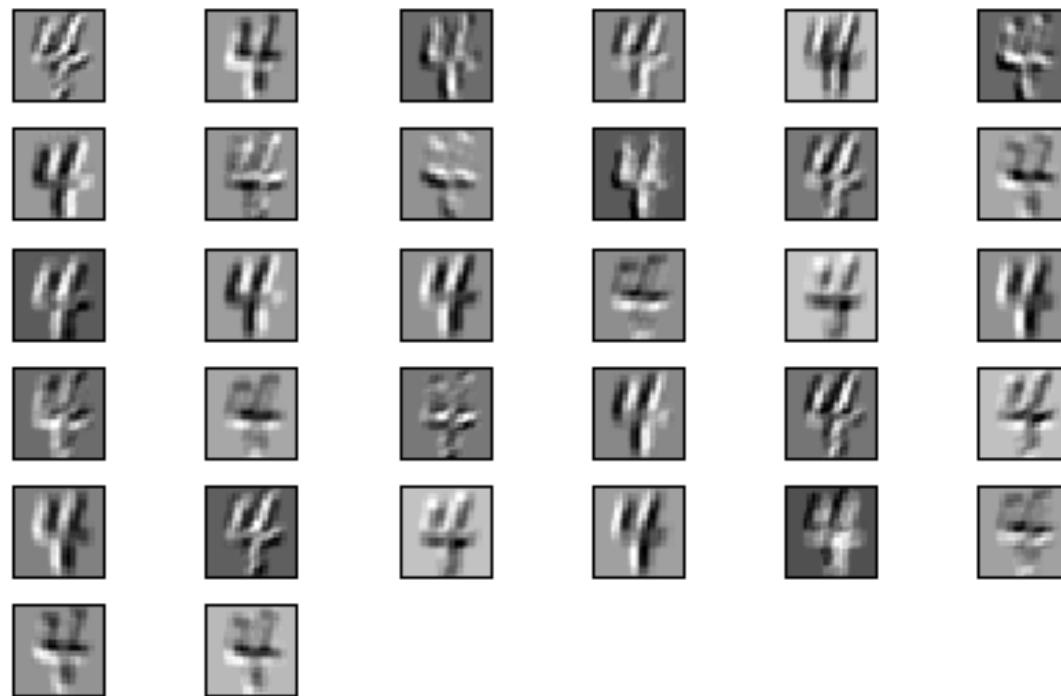
1	1	1	9	5	5	5	5
1	1	1	9	5	5	5	5
5	5	5	3	12	12	12	12
6	6	6	2	11	11	11	11
6	6	6	2	11	11	11	11

Replicate padding kích thước
(2,2)

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

Symmetric padding kích thước
(2,2)

Tầng Convolution – Feature Map



Tầng Convolution – Bài tập

- Cho ảnh RGB $5 \times 5 \times 3$ qua tầng convolutional có 2 filter $3 \times 3 \times 3$, với padding $P = 1$ và stride $S = 2$. Tính output.

Ảnh

0	0	1	1	0
0	1	1	1	0
0	1	1	1	0
0	0	1	1	0
0	0	1	1	0

Filter 1

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

Filter 2

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

0	1	1	1	0
0	1	2	2	0
0	1	1	1	0
0	1	2	2	0
0	1	1	1	0

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

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

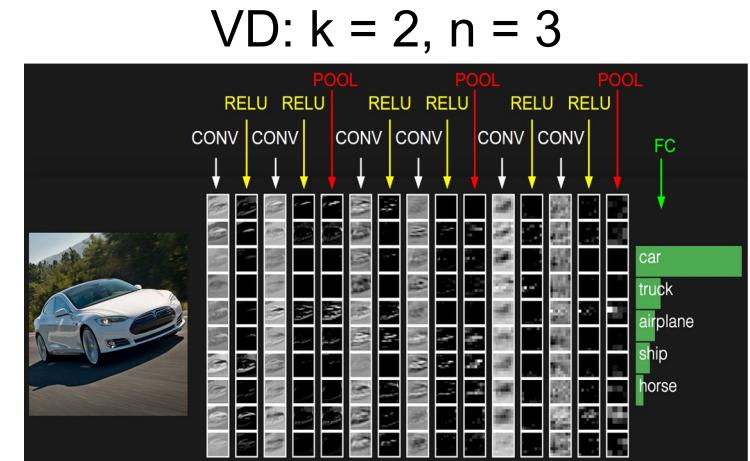
1	1	1	1	1
0	1	1	1	0
0	0	1	0	0
0	1	1	1	0
1	1	1	1	1

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

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

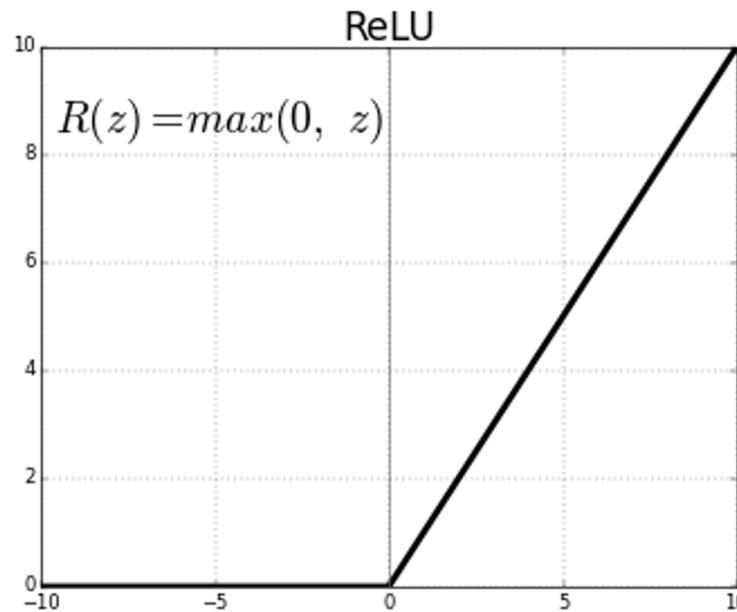
CNN - Kiến trúc phổ dụng

- Tầng convolution
 - Tầng activation (ReLU)
 - Tầng pooling (optional)
- } Lặp k lần
- Tầng fully connected
- } Lặp n lần



Tầng activation – Hàm ReLU

- Trong CNN, tầng ReLU thường theo sau tầng convolution
- Ngoài ReLU, ta có thể dùng các hàm khác: sigmoid, tanh,...



Tầng activation – Bài tập

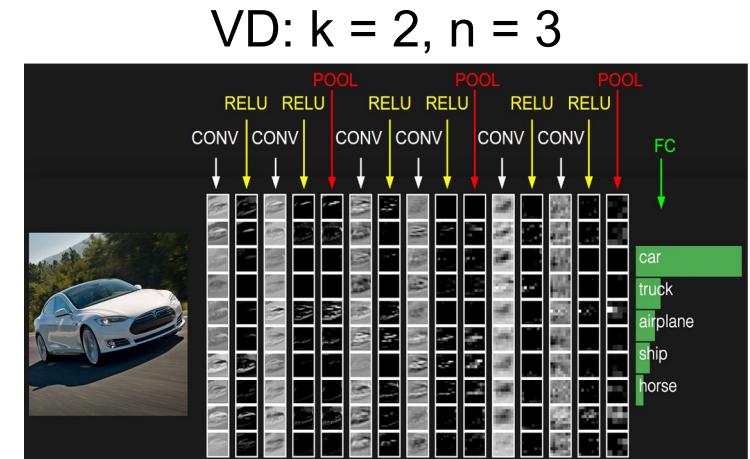
- Cho input $3 \times 3 \times 2$ như hình bên dưới. Tính output khi áp dụng tầng activation ReLU

0	-1	0
-1	-1	1
0	-2	0

1	-1	-1
0	-2	-1
0	0	0

CNN - Kiến trúc phổ dụng

- Tầng convolution
 - Tầng activation (ReLU)
 - Tầng pooling (optional)
- } Lặp k lần
- Tầng fully connected
- } Lặp n lần



Tầng Pooling - Max pooling, average pooling

- Stride thường có kích thước bằng với kích thước filter
- Filter được áp dụng độc lập trên chiều độ sâu
- Giảm chiều feature map, vị trí tuyệt đối của các đặc trưng không còn quan trọng

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

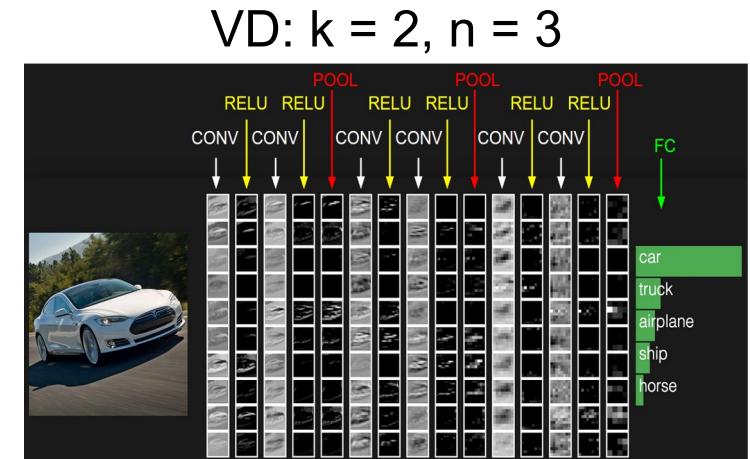
đầu vào

3	8
9	7

sau khi pooling

CNN - Kiến trúc phổ dụng

- Tầng convolution
 - Tầng activation (ReLU)
 - Tầng pooling (optional)
- } Lặp k lần
- Tầng fully connected
- } Lặp n lần



Tầng Fully Connected - Flattening

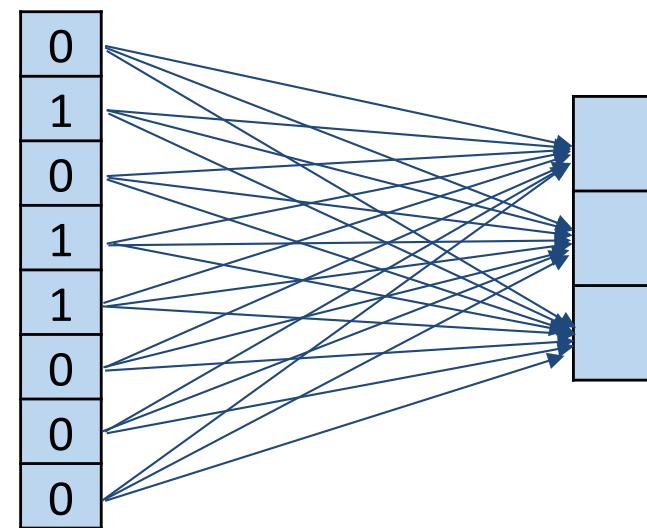
- Flattening là quá trình “duỗi” tensor 3D thành tensor 1D để làm đầu vào cho tầng fully connected
- Ví dụ:

0	1
0	1

1	0
0	0

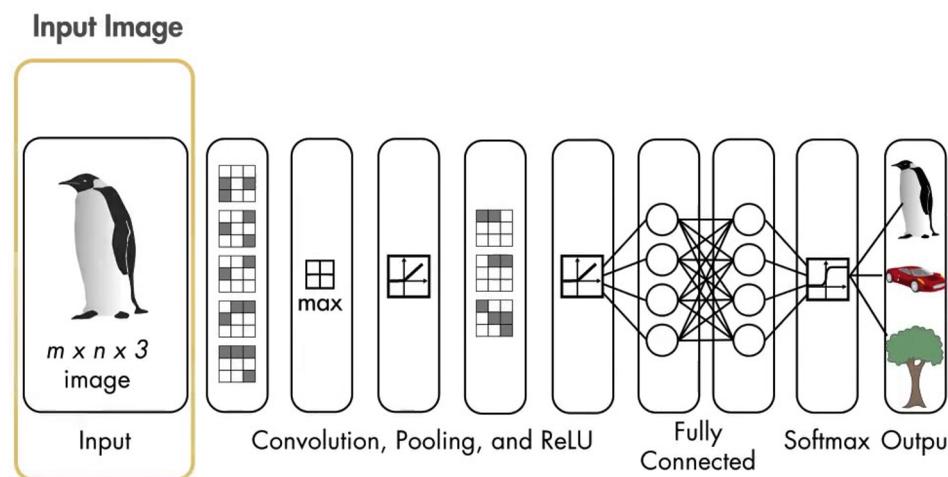
Tensor 2x2x2

Flatten



Tầng Fully Connected - Bài tập

- Thiết kế cấu trúc CNN cho bài toán phân loại 3 lớp với dữ liệu ảnh $78 \times 78 \times 3$



Tổng kết

- Kiến trúc mạng CNN kế thừa từ mạng Neural Network:
 - Chia sẻ trọng số
 - Kết nối cục bộ
- Mạng CNN gồm các tầng biến đổi
 - Tầng convolution
 - Tầng activation
 - Tầng pooling
 - Tầng kết nối đầy đủ

Tổng kết

- Ứng dụng

