



Bài 10: Một số kiến trúc Convolutional Neural Network và ứng dụng

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

- Mạng CNN - Kiến trúc thông thường
- Các ứng dụng của CNN
- Sử dụng các mạng CNN đã huấn luyện sẵn



Convolutional Neural Network

Mạng CNN - Kiến trúc thông thường

- Một CNN cho bài toán phân loại thường có cấu trúc như sau:
 - Input
 - -> Conv1_1 -> ReLU -> Conv1_2 -> ReLU (Stack 1)
 - -> Pooling1
 - -> Conv2_1 -> ReLU -> Conv2_2 -> ReLU (Stack 2)
 - -> Pooling2
 - ...
 - -> FC (Một hoặc nhiều lớp FC liên tiếp)
- Hầu như tất cả các CNN hiện nay đều dùng ReLU, nếu như trong bài báo không nhắc đến hàm kích hoạt nghĩa là họ dùng ReLU.
- Các tầng convolution giữa các lớp pooling thường được gom lại thành một nhóm và gọi là convolution stack.



Convolutional Neural Network

Các ứng dụng của CNN

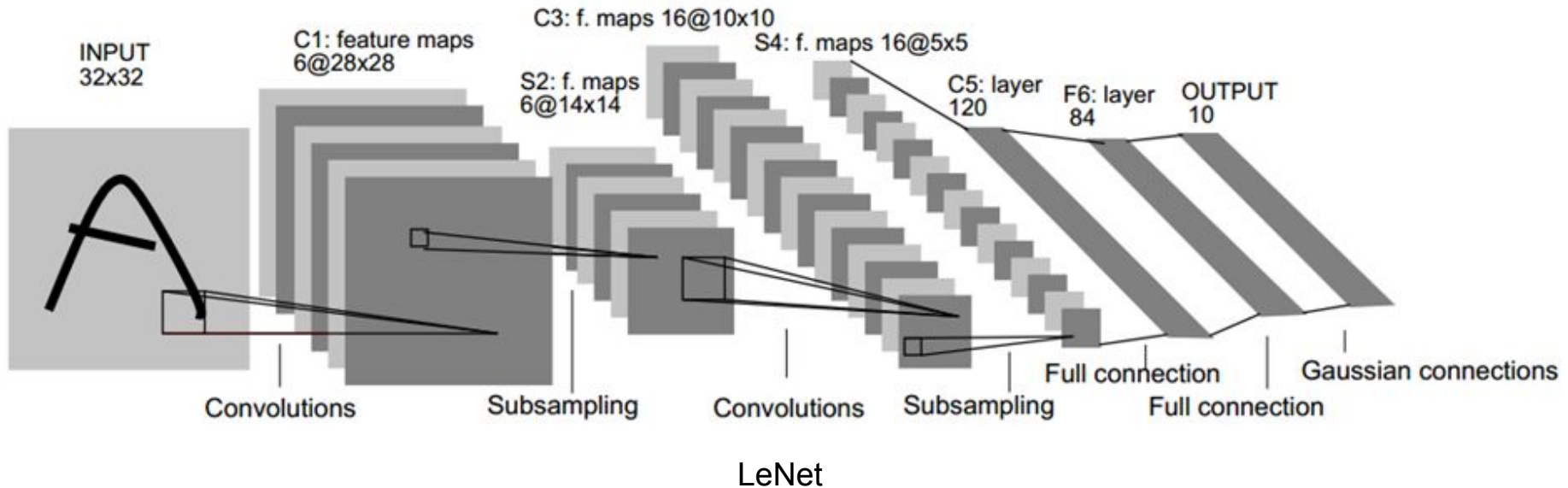
CNN có rất nhiều ứng dụng:

- Phân loại đối tượng trong ảnh (image classification)
- Phát hiện đối tượng (object detection)
- Phân đoạn ảnh (image segmentation)
- Kết hợp với recurrent neural network: image caption, visual question answering
- Xử lý ảnh



Convolutional Neural Network

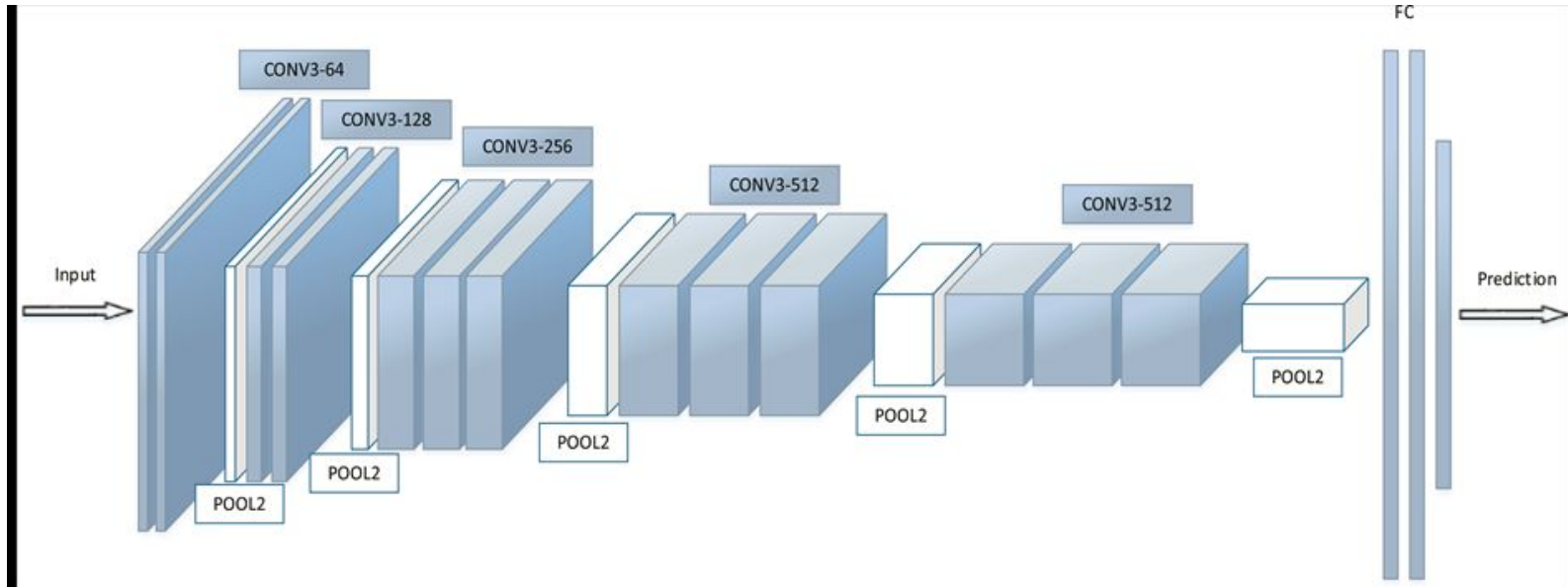
Một số CNN phổ biến cho bài toán phân lớp





Convolutional Neural Network

Một số CNN phổ biến cho bài toán phân lớp

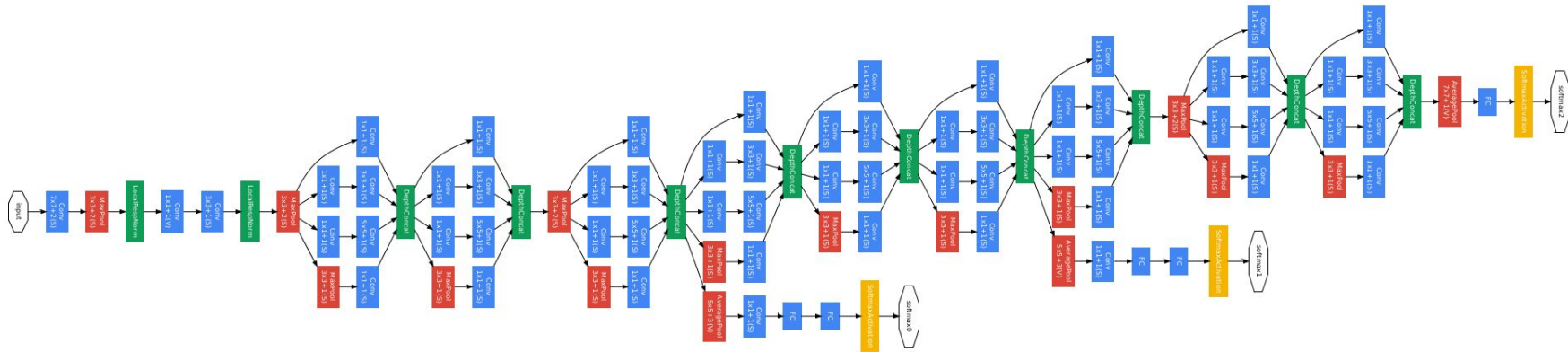


VGG-16 (Visual Geometry Group)



Convolutional Neural Network

Một số CNN phổ biến cho bài toán phân lớp



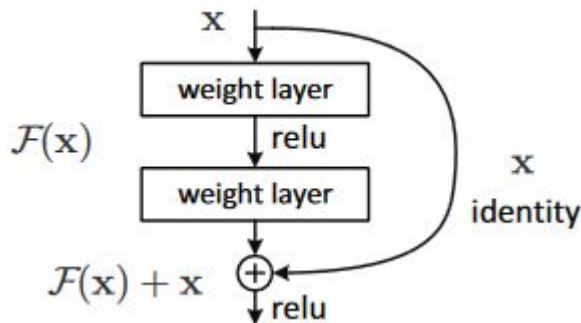
GoogleNet (Inception)



Convolutional Neural Network

Một số CNN phổ biến cho bài toán phân lớp

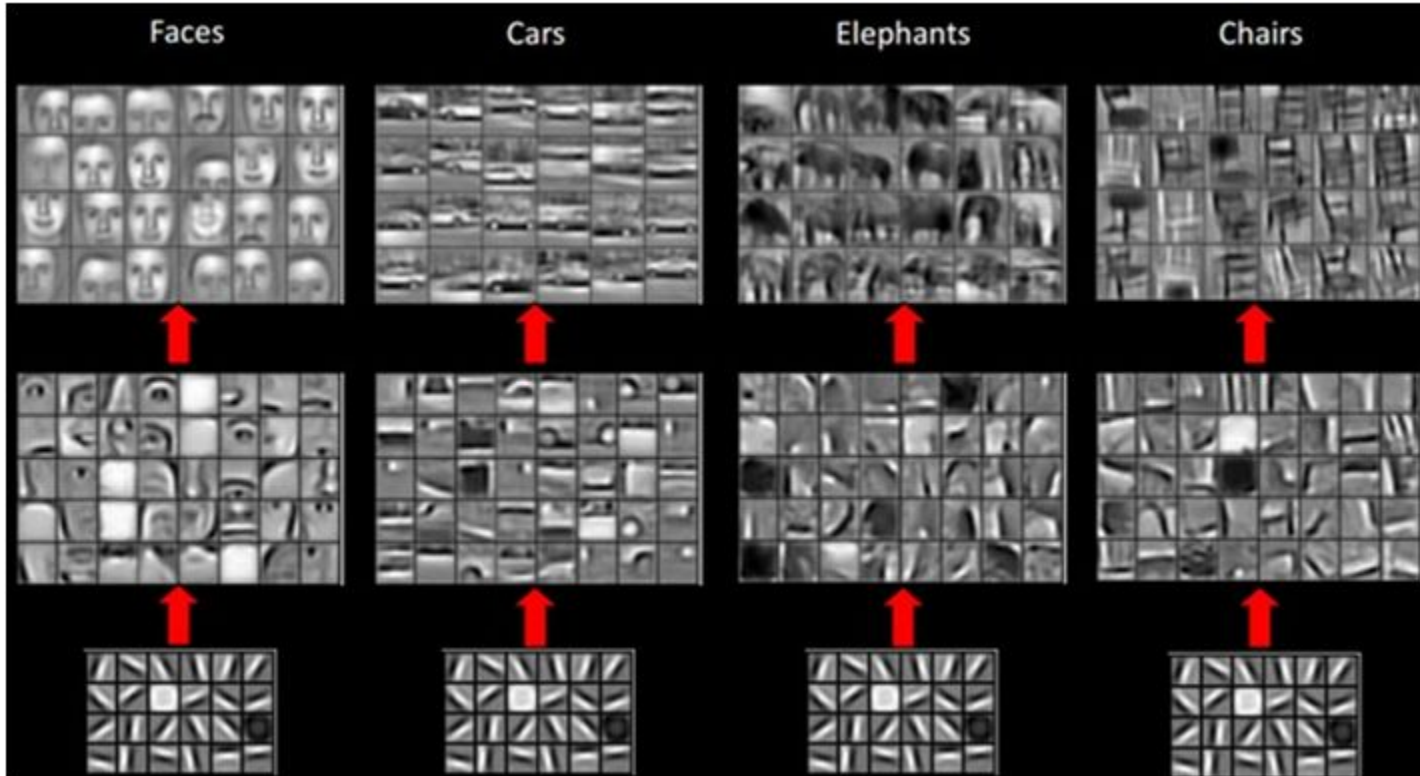
- ResNet là một trong những mạng CNN phổ biến nhất.
- Ngoài các tầng convolution thông thường, trong ResNet còn có thêm một số kết nối identity: cộng input đầu vào của tầng convolution (x) với output ($F(x)$).





Convolutional Neural Network

Rút trích đặc trưng trong CNN





Convolutional Neural Network

Các ứng dụng khác của CNN - Phát hiện vật thể (object detection)

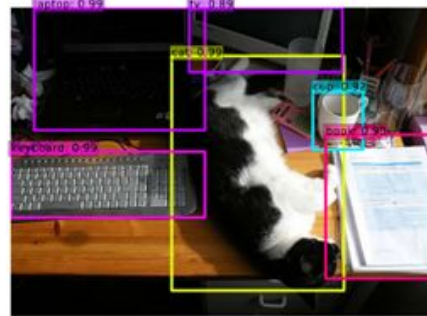
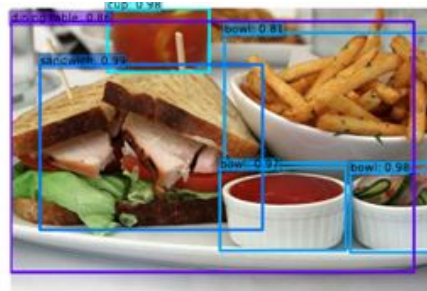
▪ Mã hóa output của CNN để nó có thể biểu diễn được các khung bao:

- Trong bài toán phân lớp, output là: [Class 1 score, Class 2 score, etc.]
- Trong bài toán phát hiện, output sẽ là: [Class 1 score, Class 2 score, ..., x, y, w, h]
- Để phát hiện nhiều đối tượng khác nhau, output sẽ có dạng:

$$\begin{bmatrix} scores_1 & x_1 & y_1 & w_1 & h_1 \\ scores_2 & x_2 & y_2 & w_2 & h_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ scores_{n-1} & x_{n-1} & y_{n-1} & w_{n-1} & h_{n-1} \\ scores_n & x_n & y_n & w_n & h_n \end{bmatrix}$$

⇒ Ở tầng convolution cuối, ta cần có $n * (\text{num_class} + 4)$ bộ kernel để sinh ra được output như trên.

▪ Một số mạng CNN dành cho việc phát hiện: SSD, YOLO





Convolutional Neural Network

Các ứng dụng khác của CNN - Phân đoạn vật thể (object segmentation)

- Bài toán phân đoạn ảnh là bài toán gán nhãn cho từng pixel của ảnh đầu vào xem pixel đó thuộc loại đối tượng nào.
- Output của mạng CNN này là một map có kích thước H và W bằng với ảnh đầu vào.
- Output map này có số layer bằng với số lớp (class) các đối tượng ta cần phân đoạn.





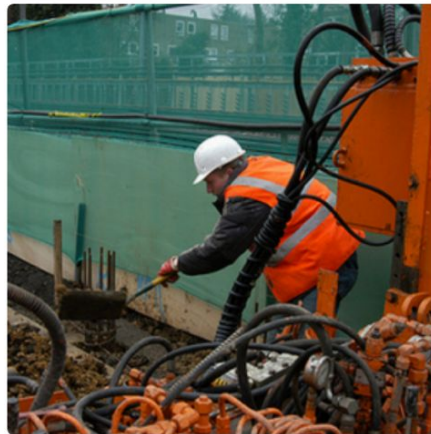
Convolutional Neural Network

Các ứng dụng khác của CNN - Image captioning

- Đặc trưng của các lớp cuối của CNN có thể được sử dụng làm input đầu vào cho mạng RNN để sinh ra câu mô tả cho ảnh.



"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



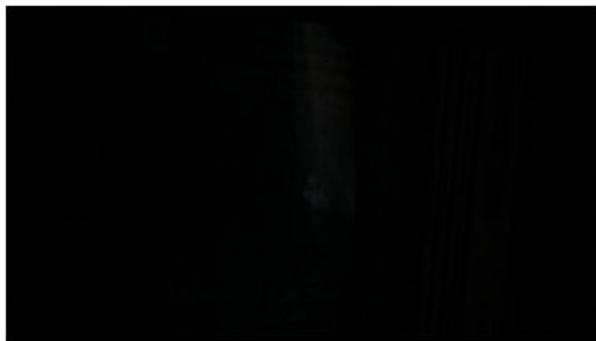
"two young girls are playing with lego toy."



Convolutional Neural Network

Các ứng dụng khác của CNN - Xử lý ảnh

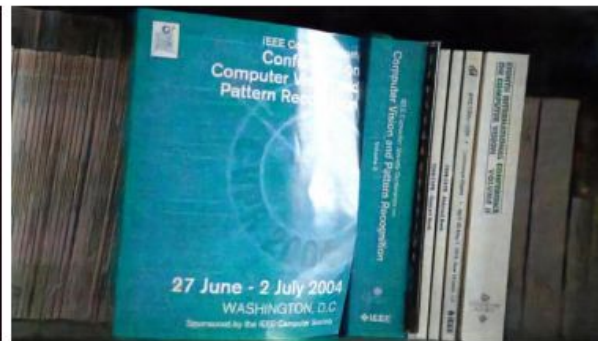
- Mạng CNN có thể được sử dụng để xử lý ảnh, nâng cao chất lượng hình ảnh, v.v



(a) Camera output with ISO 8,000



(b) Camera output with ISO 409,600



(c) Our result from the raw data of (a)



Convolutional Neural Network

Sử dụng lại các mạng CNN đã được huấn luyện sẵn

- Việc huấn luyện một mạng CNN từ đầu thường tốn rất nhiều tài nguyên, vì vậy nếu có thể ta nên sử dụng lại một số mạng CNN đã được huấn luyện sẵn trên các tập ảnh huấn luyện lớn để rút trích đặc trưng.
- Một số CNN được huấn luyện sẵn hay được dùng:
 - VGG 16 (16 layers)
 - ResNet (152 layers)
 - GoogLeNet (22 layers)
- Việc sử dụng lại một mạng NN đã huấn luyện vào một bài toán khác được gọi là transfer learning.

Convolutional Neural Network



- Ta nên sử dụng CNN trên loại dữ liệu nào?
 - Ảnh
 - Sóng âm
 - Các loại dữ liệu mà đặc trưng đầu vào có tương quan về mặt không gian
- Không nên sử dụng CNN cho các loại dữ liệu mà đặc trưng không liên quan nhau về mặt không gian.