

Concept cơ bản về CNN, ResNet, VGG

1. Convolutional Neural Networks (CNN)

1.1. Khái niệm cơ bản

CNN hoạt động dựa trên nguyên tắc sử dụng các bộ lọc (kernels) để tự động trích xuất đặc trưng từ ảnh đầu vào.

1.2. Thành phần chính của CNN

- **Convolutional Layer (Lớp tích chập):**
 - Áp dụng các bộ lọc (kernels) lên ảnh đầu vào để trích xuất đặc trưng.
 - Mỗi bộ lọc phát hiện một đặc trưng như cạnh, góc, họa tiết,...
- **Activation Function (ReLU, Tanh, Sigmoid):**
 - Tạo phi tuyến tính giúp mạng học được các đặc trưng phức tạp.
- **Pooling Layer (Lớp giảm chiều):**
 - Giảm kích thước đầu ra bằng Max Pooling hoặc Average Pooling, giúp giảm số lượng tham số và tính toán.
- **Fully Connected Layer:**
 - Kết nối tất cả các đặc trưng đã trích xuất và thực hiện phân loại.

1.3. Use case

- **Phân loại ảnh (Image Classification):** ResNet, VGG thường được dùng trong ImageNet Challenge.
- **Nhận diện đối tượng (Object Detection):** Faster R-CNN, YOLO, SSD dùng trong giám sát an ninh, xe tự hành.
- **Phân đoạn ảnh (Image Segmentation):** U-Net dùng trong y tế (phát hiện khối u, tổn thương).
- **Nhận diện khuôn mặt (Face Recognition):** Dùng trong bảo mật (Face ID của Apple).

2. VGG (Visual Geometry Group)

2.1. Concept cơ bản

VGG (VGG-16, VGG-19) là một kiến trúc CNN sâu được đề xuất bởi nhóm nghiên cứu của Oxford. Đặc điểm chính:

- **Cấu trúc đơn giản nhưng mạnh mẽ:** Chỉ sử dụng các lớp Conv 3×3 và Max Pooling 2×2 .
- **Tăng độ sâu để cải thiện hiệu suất:** VGG-16 có 16 lớp, VGG-19 có 19 lớp.
- **Nặng:** Số lượng tham số rất lớn (~138 triệu), gây tốn tài nguyên tính toán.

2.2. Use case

- **Phân loại ảnh:** VGG được sử dụng rộng rãi trong các bài toán phân loại (đạt top trong ImageNet 2014).
- **Trích xuất đặc trưng:** Dùng làm backbone cho các mô hình như Faster R-CNN.
- **Ứng dụng trong Deepfake, Style Transfer:** Vì khả năng trích xuất đặc trưng mạnh.

3. ResNet (Residual Network)

3.1. Concept cơ bản

ResNet (ResNet-18, ResNet-34, ResNet-50, ResNet-101,...) được phát triển để giải quyết vấn đề **vanishing gradient** khi mạng trở nên quá sâu.

- **Sử dụng Residual Block (Khối dư):**
 - Thay vì học trực tiếp ánh xạ $f(x)$, ResNet học phần dư $f(x) - x$, giúp gradient lan truyền dễ dàng.
 - Công thức: $y = F(x) + x$
- **Có thể đi rất sâu (hơn 100 lớp) mà vẫn train ổn định.**

3.2. Use case

- **Phân loại ảnh:** ResNet đạt top-1 trên ImageNet 2015.

- **Nhận diện đối tượng:** ResNet-50, ResNet-101 thường được dùng làm backbone cho Faster R-CNN, YOLO.
- **Ứng dụng trong y tế:** Dùng để phát hiện ung thư từ ảnh y khoa (X-ray, MRI).