

BÁO CÁO

1. Giới thiệu

ResNet, viết tắt của Residual Network, là một mạng nơ-ron tích chập sâu với đặc trưng là các kết nối residual giúp khắc phục vấn đề gradient biến mất trong các mạng sâu. ResNet bao gồm hai loại block chính:

- **Basic Block:** Sử dụng cho ResNet18 và ResNet34.
- **Bottleneck Block:** Sử dụng cho các mạng ResNet sâu hơn như ResNet50, ResNet101 và ResNet152.

2. Cài đặt

a. Class Block và Bottleneck

Cấu trúc của mã ResNet trong phần đầu của code bao gồm các thành phần sau:

- **Class Block:** Đây là basic residual block, với $\text{expansion}=1$, có hai lớp tích chập (Conv) 3×3 và Batch Normalization.
- **Class Bottleneck:** Đây là bottleneck residual block, với $\text{expansion}=4$. Nó bao gồm ba lớp tích chập 1×1 , 3×3 , và 1×1 , trong đó lớp Conv 1×1 đầu tiên giúp giảm chiều và lớp Conv 1×1 cuối cùng mở rộng chiều.

Cả Block và Bottleneck đều có một shortcut (residual connection), giúp đầu vào và đầu ra có thể được cộng trực tiếp với nhau nếu có cùng kích thước. Nếu không, một lớp downsample (Conv 1×1 với stride phù hợp) sẽ được áp dụng để đảm bảo kích thước phù hợp.

b. Lớp ResNet

Lớp ResNet được thiết kế để chứa các lớp tích chập và các block residual theo cấu hình cho từng loại ResNet:

- **Các thành phần chính của ResNet:**
 - **conv1:** Lớp tích chập đầu tiên có kernel 7×7 và $\text{stride}=2$, giúp giảm kích thước nhanh chóng.
 - **max_pool:** Pooling để giảm tiếp kích thước của đầu vào.
 - **Các Layer Residual** (layer1, layer2, layer3, layer4): Mỗi layer gồm một số lượng Block hoặc Bottleneck khác nhau, với các cấu hình tùy thuộc vào kiến trúc cụ thể.

- **avgpool** và **fc**: AdaptiveAvgPool2d giúp cố định đầu ra của mạng trước lớp fully connected (FC), và lớp FC chuyển đổi đầu ra thành vector có độ dài tương ứng với số lớp (số nhãn phân loại).

c. Hàm `_make_layer`

Hàm `_make_layer` chịu trách nhiệm tạo các layer residual gồm nhiều block residual (hoặc bottleneck), với các chi tiết:

- **Downsampling**: Khi `stride != 1` hoặc số kênh đầu vào khác số kênh đầu ra, một lớp Conv1x1 sẽ được dùng để giảm kích thước (downsample).
- **Lặp block**: Sau khi tạo block đầu tiên (có thể có downsampling), các block tiếp theo sẽ có `stride=1`.