

Báo cáo tóm tắt kết quả thí nghiệm phân loại ảnh CIFAR-10

1. Đề bài

Phân loại ảnh CIFAR-10 sử dụng các kỹ thuật deep learning. CIFAR-10 là bộ dữ liệu gồm 10 lớp đối tượng: máy bay, ô tô, chim, mèo, hươu, chó, ếch, ngựa, tàu thủy và xe tải.

2. Tiền xử lý dữ liệu

- **Phân chia dữ liệu:**
 - Tập huấn luyện: 90% của tập train gốc
 - Tập validation: 10% của tập train gốc
 - Tập kiểm tra: Sử dụng tập test gốc của CIFAR-10
- **Kỹ thuật tăng cường dữ liệu (Data Augmentation):**
 - Lật ngang ảnh ngẫu nhiên (Random horizontal flip)
 - Cắt ngẫu nhiên kích thước 32x32 với padding 4 pixels
- **Chuẩn hóa dữ liệu:**
 - Giá trị trung bình: (0.4914, 0.4822, 0.4465)
 - Độ lệch chuẩn: (0.2470, 0.2435, 0.2616)

3. Thiết kế mô hình

BasicCNN

Mạng CNN đơn giản với cấu trúc:

- 3 khối tích chập, mỗi khối gồm:
 - Lớp tích chập 2D (Conv2D)
 - Chuẩn hóa theo batch (BatchNorm)
 - Hàm kích hoạt ReLU
 - Lớp gộp cực đại (MaxPool)
- 2 lớp fully connected với dropout để giảm overfitting

CustomEfficientNet

Mô hình dựa trên kiến trúc EfficientNet-B0 với các điều chỉnh:

- Sử dụng pre-trained weights từ ImageNet
- Điều chỉnh để phù hợp với kích thước ảnh 32x32 của CIFAR-10
- Thay đổi lớp phân loại cuối cùng để phù hợp với 10 lớp của CIFAR-10

4. Siêu tham số và cấu hình

Các thí nghiệm được thực hiện với nhiều cấu hình khác nhau:

BasicCNN:

- 3 cấu hình khác nhau:
 1. Batch size=256, lr=0.005, weight_decay=5e-4, optimizer=adam, scheduler=cosine, epochs=10
 2. Batch size=512, lr=0.01, weight_decay=1e-4, optimizer=sgd, scheduler=step, epochs=10
 3. Batch size=128, lr=0.001, weight_decay=1e-3, optimizer=adam, scheduler=none, epochs=10

CustomEfficientNet:

- 3 cấu hình khác nhau:
 1. Batch size=256, lr=0.005, weight_decay=1e-4, optimizer=adam, scheduler=cosine, epochs=5
 2. Batch size=512, lr=0.01, weight_decay=5e-4, optimizer=sgd, scheduler=step, epochs=5
 3. Batch size=128, lr=0.001, weight_decay=1e-3, optimizer=adam, scheduler=none, epochs=5

5. Kết quả

Độ chính xác trung bình

- **BasicCNN:** 74.84% \pm 1.39%
- **CustomEfficientNet:** 80.23% \pm 3.70%

Mô hình tốt nhất

- **Loại mô hình:** CustomEfficientNet

- **Cấu hình:** Batch size=256, lr=0.005, weight_decay=1e-4, optimizer=adam, scheduler=cosine, epochs=5
- **Độ chính xác trên tập kiểm tra:** 84.59%
- **Loss trên tập kiểm tra:** 0.4492

Chi tiết kết quả theo từng mô hình

1. **BasicCNN** (tốt nhất): 76.29% (Adam, lr=0.001, batch_size=128)
2. **CustomEfficientNet** (tốt nhất): 84.59% (Adam, lr=0.005, batch_size=256)

6. Nhận xét

1. **CustomEfficientNet hiệu quả hơn BasicCNN:** Mô hình dựa trên kiến trúc EfficientNet cho kết quả tốt hơn đáng kể so với mô hình CNN cơ bản, chứng tỏ lợi ích của việc sử dụng mô hình được thiết kế tối ưu và pre-trained.
2. **Optimizer Adam hiệu quả hơn SGD:** Trong cả hai loại mô hình, các cấu hình sử dụng Adam optimizer đều cho kết quả tốt hơn so với SGD.
3. **Learning rate scheduler:** Lịch điều chỉnh learning rate kiểu cosine annealing cho kết quả tốt nhất, giúp mô hình hội tụ nhanh và ổn định hơn.
4. **Độ biến thiên kết quả:** CustomEfficientNet có độ lệch chuẩn cao hơn (3.70% so với 1.39% của BasicCNN), cho thấy mô hình này nhạy cảm hơn với sự thay đổi của hyperparameters.

Nhìn chung, thí nghiệm cho thấy mô hình CustomEfficientNet với Adam optimizer và lịch cosine annealing là lựa chọn tốt nhất cho bài toán phân loại ảnh CIFAR-10, đạt độ chính xác 84.59% trên tập kiểm tra.