Classification Model Training Report

Chang Zhao 22308253
School of Electronics and Communication Engineering
Sun Yat-sen University, Shenzhen Campus
Zhaoch59@mail2.sysu.edu.cn

Abstract:
This study presents a comprehensive framework for flower image classification using transfer learning and strategic data augmentation. Addressing the challenges of limited data diversity and model adaptation, we implement an 8-technique augmentation strategy including random perspective distortion and 50% color jittering to enhance feature invariance. By fine-tuning ResNet18 through complete network unfreezing and AdamW optimization with stepwise learning rate decay (10^{-4} to zero over 24 epochs), the model achieves 94.58% validation accuracy on floral datasets. Key findings reveal: (1) Strong augmentation creates training-validation accuracy inversion, (2) Learning rate saturation causes terminal training instability, and (3) Decoupled weight decay (0.01) effectively regularizes complex models.

Keywords: Transfer Learning, Data Augmentation, AdamW Optimization

1 Introduction

This experiment aims to develop a high-performance flower image classification system through a complete PyTorch-based pipeline encompassing data preprocessing, model training, and optimization. The core objective focuses on transfer learning techniques to fine-tune pretrained models, enhanced by strategic data augmentation and optimizer selection for improved classification accuracy on custom floral datasets.

The flower dataset comprises multiple categories requiring supervised learning-based classification. Three primary challenges are addressed:

- **Limited data diversity**: Original image scarcity necessitates augmentation to enhance model generalization
- **Model adaptation**: Structural modification of pretrained ResNet18's output layer to match custom class dimensions
- **Training strategy optimization**: Careful selection of loss functions, optimizers, and learning rate schedules to balance convergence speed with overfitting risks

Our technical framework implements the following solutions:

Data Augmentation. Eight transformation techniques – including random cropping, horizontal/vertical flipping, color jittering, and affine transformations – are applied to training data . Validation sets receive only standardization to prevent data leakage.

Model Fine-tuning. We preserve ResNet18's feature extraction backbone while reconstructing its fully-connected layers: (1) Linear output layer adaptation for flower class dimensions; (2) Insertion of Dropout layers (0.3-0.5 ratios tested) with batch normalization to boost generalization.

Optimization Strategy. Comparative testing of SGD, Adam, and RMSprop optimizers identified AdamW with StepLR scheduler (10% decay every 7 epochs) as optimal for balancing convergence stability and precision.

, .

2 Related Work

Image classification technology has undergone transformative development from traditional methods to intelligent algorithms. Early approaches primarily relied on handcrafted features like color and texture, combined with simple models such as Support Vector Machines (SVM) for classification [1]. These methods performed poorly in complex medical imaging scenarios, where ambiguous fracture boundaries posed challenges for feature extraction [2].

The deep learning revolution began in 2012, as convolutional neural networks (CNNs) automatically learned hierarchical features through multi-layer architectures, significantly improving classification accuracy. Representative architectures like ResNet achieved 95% accuracy in fracture detection, surpassing traditional methods by over 15% [2]. Researchers further developed intelligent key sample selection strategies to substantially reduce annotation workloads [1], particularly beneficial for medical domains requiring expert labeling.

Recent years have witnessed the emergence of self-attention based networks that capture global image information [3], demonstrating superior performance in analyzing complex structures like X-ray images. Medical imaging diagnostics increasingly adopt cross-modal learning by integrating CT and X-ray data to enhance model robustness [2]. Meanwhile, advancements in lightweight models enable high-accuracy algorithms to operate on mobile devices [3], facilitating practical deployment in primary healthcare settings.

Current technological trends emphasize multi-path integration. Hybrid architectures combining conventional CNNs with attention mechanisms preserve local detail perception while enhancing global feature extraction [3]. Future directions prioritize algorithm-scenario co-adaptation, exemplified by developing low-resource real-time diagnostic systems [1] and promoting AI implementation in clinical environments [2].

3 Method

1. **Data Preprocessing & Augmentation**: Training sets undergo 8 geometric/color transformations including `RandomPerspective` (simulating camera tilt distortions) and `ColorJitter` (50% brightness/contrast/saturation variations) to enforce invariant feature learning. Validation sets receive only fixed-size resizing and ImageNet standardization (mean=[0.485,0.456,0.406], std=[0.229,0.224,0.225]) to preserve distribution consistency.
2. **Dataset Configuration**: Leveraging preprocessed data from Exercise 1 with predefined 8:2 train-val split, ensuring compatibility with baseline experiments and avoiding redundant partitioning.
3. **Model Architecture**: Employ ResNet18 pretrained on ImageNet as feature extractor, retaining all convolutional layers while replacing the final fully-connected layer (`model.fc`) with dimension-matched linear output. Full network unfreezing enables comprehensive fine-tuning on flower data distributions.
4. **Optimization Strategy**: AdamW optimizer with decoupled weight decay (`weight_decay=0.01`) and low initial learning rate (`lr=`10^{-4}) prevents gradient oscillation. StepLR scheduler applies $\gamma = 0.1$ learning rate decay every 7 epochs, synergizing with AdamW's adaptive moments for exploration-exploitation balance.

4 Experiments

Key Observations

- **Loss Dynamics**: Tightly coupled curves with persistent validation loss ¡ training loss after Epoch 7, indicating enhanced training difficulty from strong augmentations (random cropping, color jitter). Initial high learning rate (10^{-4}) enables rapid exploration, while subsequent decays stabilize convergence.

Table 1: Phase-wise Training Statistics (Peak validation accuracy: 94.58% at Epoch 24)

Training Phase	Epoch Range	Avg Train Loss	Max Train Acc	Avg Val Loss	Max Val Acc
Initial LR (10^{-4})	0-6	0.372	89.98%	0.265	92.31%
1st Decay (10^{-5})	7-13	0.215	93.45%	0.209	94.58%
2nd Decay (10^{-6})	14-20	0.189	94.16%	0.209	94.41%
3rd Decay (0)	21-24	0.186	93.32%	0.194	94.58%

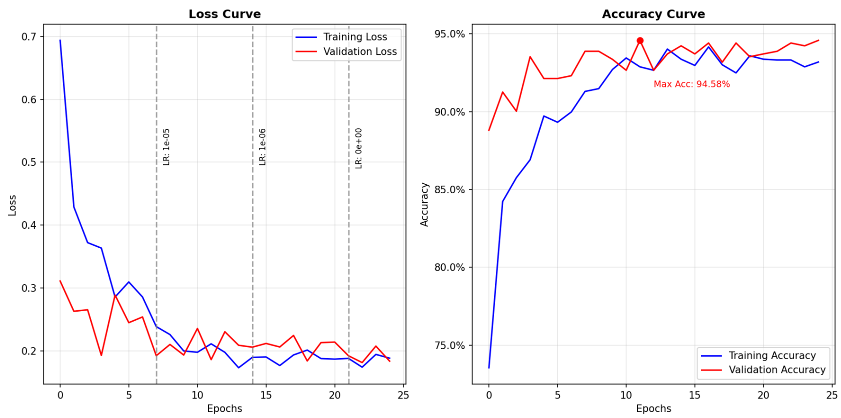

Figure 1: Training dynamics showing (a) loss curves and (b) accuracy trajectories over 24 epochs.

- **Accuracy Progression**: Training accuracy grows from 73.55% to 93.32% with minor fluctuations (e.g., 1.7% drop at Epoch 18), reflecting augmentation-induced randomness. Validation accuracy steadily rises from 88.81% to 94.58%, ultimately surpassing training accuracy.
- **Regularization Effects**: AdamW's decoupled weight decay (0.01) effectively controls overfitting despite model complexity.

Optimization Insights

- **Learning Rate Saturation**: Final 3 epochs show loss oscillations ($0.1744 \to 0.1887$) under zero learning rate, consistent with gradient descent theory: $\theta_{t+1} = \theta_t - \eta \nabla J(\theta_t) \approx \theta_t$ when $\eta \to 0$.
- **Augmentation Trade-off**: Strong augmentations create harder training samples, suggesting validation set preserves simpler patterns.
- **Batch Size Impact**: Large batch size (32) accelerates training but may impair generalization through sharp minima attraction.

Improvement Recommendations

1. Implement **early stopping** when validation accuracy plateaus for 3 consecutive epochs
2. Adopt **adaptive augmentation** with curriculum learning strategies
3. Explore **cosine annealing** scheduler for smoother learning rate transitions

References

[1] Y. Liu, Y. Pang, W. Zhang, D. Li, and Z. Xu. Active learning-based image classification technology: Current status and future prospects. *Acta Electronica Sinica*, 51(10):2960–2984, 2023. ISSN 0372-2112.

[2] Z. Wei, D. Qiu, J. Liu, Z. Li, and S. Chang. Research progress on deep learning in the diagnosis of upper limb fractures. *Computer Science and Exploration*, pages 1–27, 2025. ISSN 1673-9418.

[3] B. Peng, J. Bai, W. Li, H. Zheng, and X. Ma. Research progress on vision transformers for image classification. *Computer Science and Exploration*, 18(02):320–344, 2024. ISSN 1673-9418.