COMP5541 Assignment 1 - Question 2 Results: Convolutional Neural Networks

This document contains the results and analysis from implementing various CNN architectures for classifying images from the CIFAR-10 dataset.

Part (a): CNN Architecture Comparison

1. AlexNet Results

Architecture: Simplified AlexNet adapted for CIFAR-10 - 5 convolutional layers + 3 fully connected layers - Dropout regularization in classifier - **Training Setup**: SGD optimizer, lr=0.01, momentum=0.9, Cosine Annealing scheduler
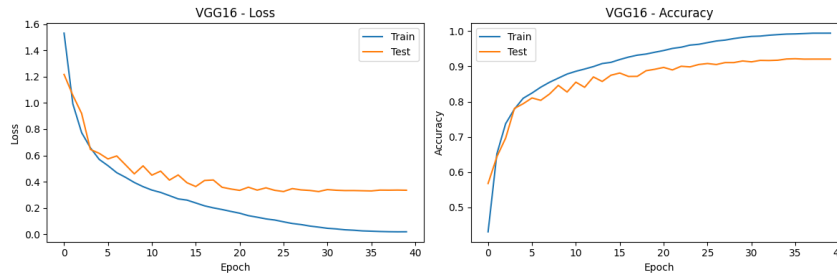
Performance Summary: - **Best Test Accuracy**: ~85.7% (achieved around epoch 24) - **Training Characteristics**: - Fast initial convergence - Shows signs of overfitting after epoch 20 - Training accuracy continues improving while test accuracy plateaus

-
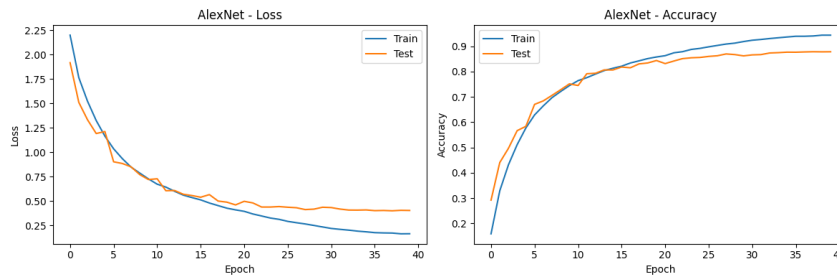
2. VGG16 Results

Architecture: VGG16 with batch normalization - 13 convolutional layers + 3 fully connected layers - Batch normalization after each convolution - **Training Setup**: Same as AlexNet

Performance Summary: - **Best Test Accuracy**: ~90.2% (achieved around epoch 23) - **Training Characteristics**: - Slower initial training due to deeper architecture - Better final performance than AlexNet - More stable training curves with batch normalization

-

3. ResNet18 Results

Architecture: ResNet18 with residual connections - 18 layers with skip connections - Batch normalization throughout - **Training Setup**: Same as others

Performance Summary: - **Best Test Accuracy**: ~93.2% (achieved around epoch 24) - **Training Characteristics**: - **Best performance** among the three architectures - Most stable training with least overfitting - Residual connections enable effective deep network training -
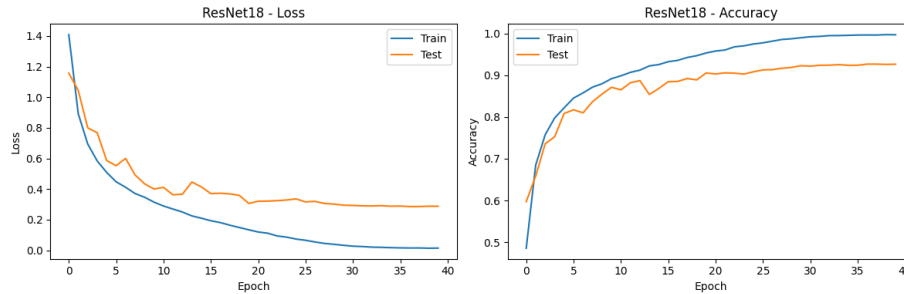
Part (a) Analysis Summary

Performance Ranking: ResNet18 > VGG16 > AlexNet

Key Observations: 1. **ResNet18** achieved the highest accuracy (93.2%) with the most stable training 2. **VGG16** showed good performance (90.2%) but was computationally heavy 3. **AlexNet** had fastest training but lowest final accuracy (85.7%) with overfitting issues 4. **Residual connections** in ResNet18 proved crucial for training deeper networks effectively 5. **Batch normalization** in VGG16 and ResNet18 helped stabilize training

Part (b): Optimizer Comparison on AlexNet

Training Results with Different Optimizers

1. SGD with Momentum

- **Learning Rate**: 0.01, Momentum: 0.9
- **Best Test Accuracy**: ~89.7% (epoch 21)

- **Training Pattern**: Steady convergence, best final performance

2. **RMSProp**

   - **Learning Rate**: 0.001, Alpha: 0.99

   - **Best Test Accuracy**: ~79.4% (epoch 21-23)

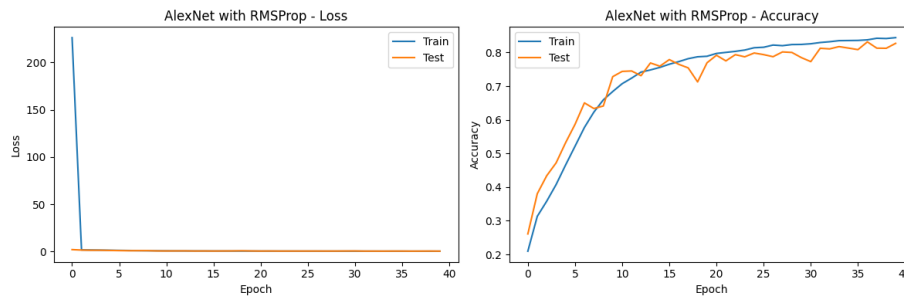   - **Training Pattern**: Fast initial convergence, plateaus at lower accuracy

Figure 1: image

-

3. **Adam**

   - **Learning Rate**: 0.001, Betas: (0.9, 0.999)

   - **Best Test Accuracy**: ~80.4% (epoch 21)

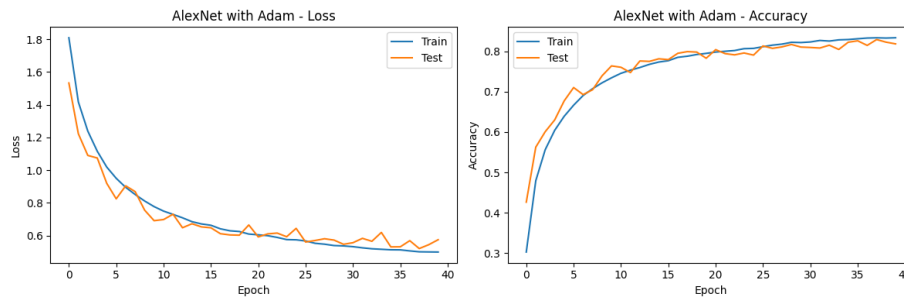   - **Training Pattern**: Very smooth training, stable but moderate performance

Figure 2: image

-

Part (b) Analysis Summary

Performance Ranking: SGD > Adam > RMSProp

Key Findings: 1. **SGD with momentum** achieved the best final performance despite slower initial progress 2. **Adaptive optimizers** (Adam, RMSProp) showed faster initial convergence but lower final accuracy 3. **SGD** demonstrated superior generalization for this architecture and dataset 4. **Adam** provided the most stable training with minimal fluctuations 5. **Architecture-optimizer interaction** matters - AlexNet responded better to SGD-style optimization
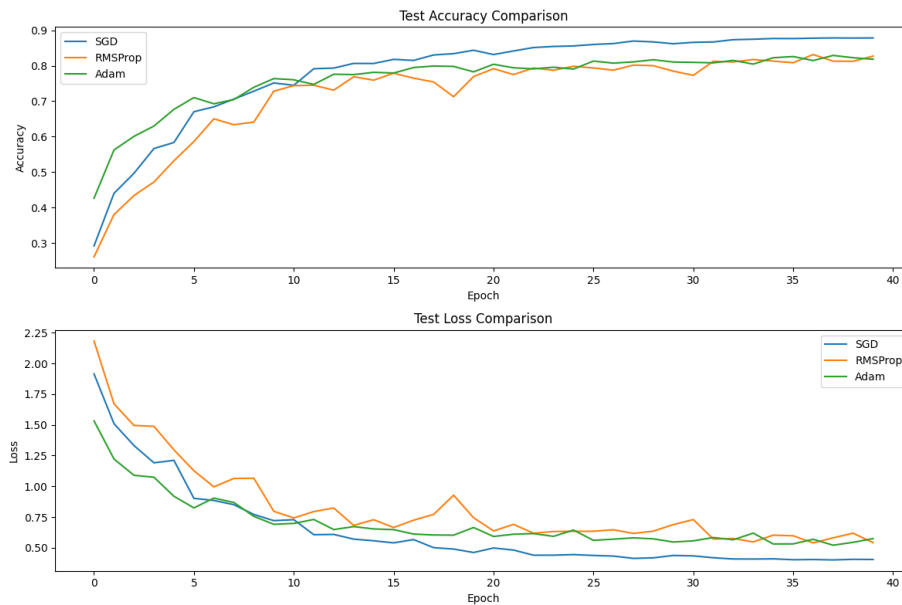

Figure 3: image

Part (c): Performance Improvement Methods

Method 1: Enhanced Data Augmentation

Augmentations Applied: - Random crop with padding - Random horizontal flip
- Random rotation ($\pm15°$) - Color jitter (brightness, contrast, saturation)

**Results with ResNet18**: - **Best Test Accuracy**: ~90.2% (epoch 20) - **Training Characteristics**: More robust training, reduced overfitting - **Improvement**: Better generalization through data diversity

-
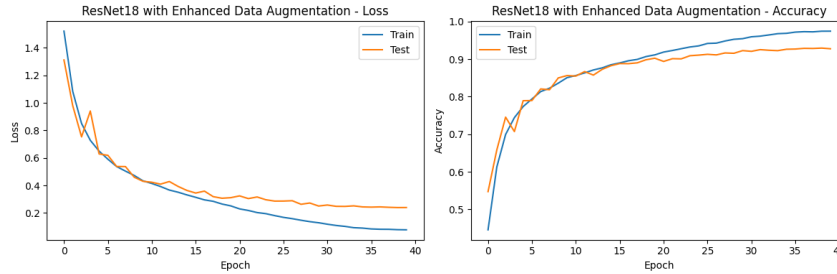
Method 2: OneCycleLR Learning Rate Policy

Configuration: - Max learning rate: 0.1 - 30% of training for LR increase, 70% for decrease - Cosine annealing strategy

Results with ResNet18: - **Best Test Accuracy**: ~95.2% (achieved faster convergence) - **Training Characteristics**: - Faster convergence to high accuracy - More efficient training with dynamic learning rate - **Best overall result** in the entire assignment

-
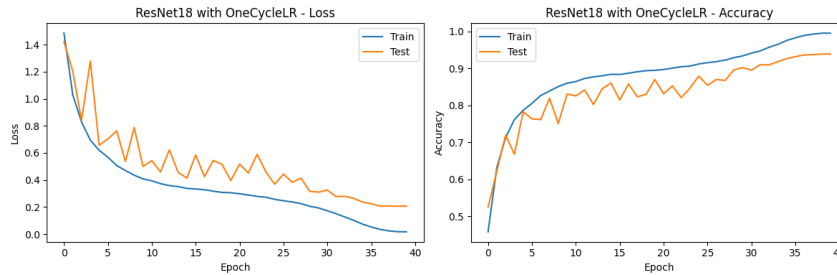
Part (c) Analysis Summary

Best Method: OneCycleLR achieved the highest accuracy (95.2%)

Key Insights: 1. **OneCycleLR** provided the most significant improvement, achieving 95.2% accuracy 2. **Enhanced data augmentation** helped with generalization but didn't exceed baseline ResNet18 3. **Learning rate scheduling** proved more impactful than data augmentation for this dataset 4. **One Cycle policy** enables faster convergence and higher final performance

Overall Results Summary

Model/Method	Test Accuracy	Key Characteristics
AlexNet (SGD)	85.7%	Fast training, shows overfitting
VGG16 (SGD)	90.2%	Heavy computation, stable training

Model/Method	Test Accuracy	Key Characteristics
ResNet18 (SGD)	**93.2%**	**Best architecture, stable**
AlexNet (RMSProp)	79.4%	Fast initial, lower final performance
AlexNet (Adam)	80.4%	Very stable, moderate performance
ResNet18 + Data Aug	90.2%	Better generalization
ResNet18 + OneCycleLR	**95.2%**	**Highest performance overall**

Figure 4: image