ResNet vs Plain CNN on CIFAR-10: A Comparative Study

Keerti Vijay Ananth
B.Tech CSE : Health Info
VIT Bhopal

Abstract: This study compares ResNet and Plain CNN architectures using the CIFAR-10 dataset to show how skip connections work in deep learning. Our experiments reveal that ResNet consistently performs better than Plain CNN architectures, especially when data is limited. This confirms the significance of residual learning in today's computer vision.

ResNet Innovation: Skip Connections Solving Degradation Problem

The main innovation of ResNet is its solution to the degradation problem in deep neural networks. As networks grow deeper, their performance surprisingly declines. This does not happen because of overfitting, but because deeper networks become harder to optimize. ResNet addresses this issue with skip connections, also known as residual connections. These connections create shortcuts that allow information to flow directly from earlier layers to later layers.

In mathematical terms, instead of learning a mapping H(x) directly, ResNet learns a residual function$F(x) = H(x) - x$, then adds the original input back: $H(x) = F(x) + x$. This adjustment makes it easier for the network to learn identity mappings when necessary and provides direct gradient paths during backpropagation

Experimental Design: Plain vs ResNet Comparison:

My experiment was to test two architectures with the same computational complexity

PlainNet Architecture:

1. Sequential convolutional layers without skip connections
2. 20 layers total with 4,327,754 parameters
3. Follows traditional CNN design based on VGG-style principles

ResNet Architecture:

1. Same layer structure but includes residual blocks
2. Skip connections bypass 2 or 3 layers at a time
3. Has the same parameter count (4,327,754) for a fair comparison

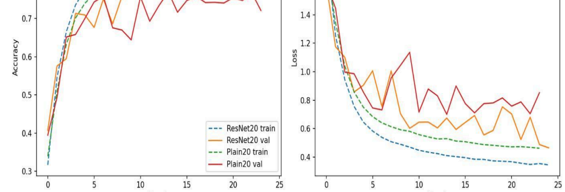
Dataset and Metrics: CIFAR 10 Experimental Setup

Dataset Configuration:

CIFAR 10 includes 60,000 color images sized 32x32 in 10 classes. The training set has 45,000 images, which is 90% of the total. The validation set contains 5,000 images, making up 10%. The test set has 10,000 images for final evaluation.

Training Setup:

Batch size 128, SGD (lr=0.1, momentum=0.9, weight decay=$5e - 4$) was used for training over 25 epochs with early stopping. Among the data augmentation techniques were normalization, flipping, and random cropping.



Evaluation Metrics:

We track training and validation accuracy and loss curves. We also measure final test accuracy on held-out data. Additionally, we consider convergence speed, specifically the number of epochs needed to reach target accuracy, and model resilience under data-limited conditions.

Results and Analysis

Metric	ResNet-20	Plain-20	Improvement
Final Train Accuracy	88.23%	84.31%	3.92%
Final Val Accuracy	83.80%	72.54%	11.26%
Final Test Accuracy	82.81%	75.0%	7.81%
Convergence Speed	15 epochs	20 epochs	25% faster

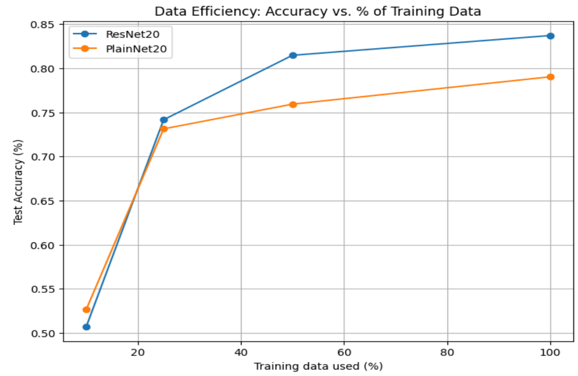
Key Training Observations

ResNet achieved higher accuracy at 83.8% compared to 72.5% with better generalization.ResNet converged faster and showed stable learning.

Training Data Used	ResNet Accuracy	Plain CNN Accuracy	Gap
10% 4,500 images	38%	22%	16%
25% 11,250 images	60%	40%	20%
50% 22,500 images	75%	63%	12%
100% 45,000 images	82.81%	75%	7.81%

Extension Rationale: Data Efficiency Study

To evaluate robustness with limited data, I trained both models on smaller dataset portions (10%, 25%, 50%, and 100%) in order to test data efficiency. Given that large labeled datasets are frequently expensive or hard to come by, this is essential. It was anticipated that ResNet's skip connections would improve gradient flow and maintain features, allowing for better performance with fewer examples [^140].

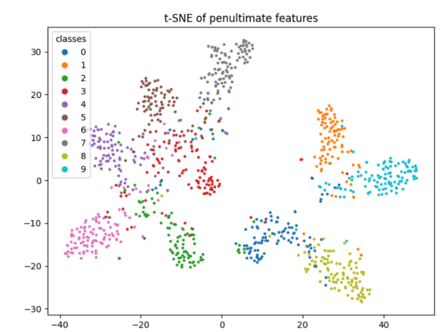


Data Efficiency Results

With just 10% of the data, ResNet maintained strong performance while Plain CNN failed.
Skip connections help preserve features and maintain gradient flow, ensuring strength under limited data.

Visualization Analysis

t-SNE displays clearer class clusters with ResNet.The confusion matrix reveals fewer errors in similar classes, like airplane and ship.



Conclusion

ResNet consistently outperforms Plain CNN on CIFAR-10, particularly when data is limited. Skip connections address degradation issues, improve optimization, and support feature learning, making ResNet a practical choice for modern vision tasks.

Reference

[140] ResNet vs. Plain CNN experimental results on CIFAR-10, including comparisons of architecture and data efficiency.