Comparative Study of CNN Architectures on CIFAR-10

1. Introduction

In this task I compared four convolutional neural network (CNN) architectures on the CIFAR-10 image classification dataset: a custom Advanced CNN, EfficientNet-B0, ResNet-18, and VGG16.The objective was to understand how these models work and get a research kind of intuitive over this more than accuracy

2. Experimental Setup

Dataset

Dataset: CIFAR-10
Classes: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck
Image size: 32×32 RGB
Split: 80% training, 20% validation, Separate 20% Testing
Augmentation: Check in the project quite a few innovative ones used 😁

Training Configuration

Model	Optimizer	Learning Rate	Weight Decay	Scheduler	Epochs
Advanced CNN	SGD (momentum=0.9)	0.01	1e-6	ReduceLROnPlateau	50
EfficientNet-B0	Adam	0.001 (most efficient)	1e-4	ReduceLROnPlateau	50
ResNet-18	SGD (momentum=0.9)	0.01	1e-4	ReduceLROnPlateau	50
VGG16	Adam	0.001	1e-4	ReduceLROnPlateau	50

3. Model Architectures

3.1 Advanced CNN (Custom Model)

A deep convolutional network with nine convolutional layers, three Batch Normalization layers, and a three-layer fully connected classifier.
Architecture Summary:

• Conv Layers: 3 → 64 → 128 → 256 → 512 at 256 and 512 we have gone deeper in extracting features as they are important

• Activation: ReLU
• Pooling: MaxPool2d (2×2)
• Normalization: BatchNorm after key convolutional blocks
• Dropout: p = 0.2
• Fully Connected Layers: [8192 → 8192 → 4096 → 10]
• Total Parameters: 108.3 million

3.2 Pretrained Architectures

EfficientNet-B0: Compound-scaled network balancing depth, width, and input resolution.
ResNet-18: Incorporates residual (skip) connections to improve gradient flow.
VGG16: Classic deep CNN with uniform 3×3 convolutions, known for robust performance.

4. Results

4.1 Overall Performance

Model	Accuracy (%)	Precision	Recall	F1-score	Train Loss	Val Loss	Val Acc (%)
Advanced CNN	91.41	0.915	0.912	0.912	0.188	0.282	91.23
EfficientNet-B0	93.03	0.869	0.867	0.868	0.378	0.399	86.74
ResNet-18	94.98	0.858	0.856	0.857	0.254	0.471	85.62
VGG16	96.92	0.915	0.915	0.915	0.153	0.305	91.52

4.2 Class-wise Performance

Class	Advanced CNN F1	EfficientNet F1	ResNet F1	VGG16 F1
Airplane	0.914	0.938	0.962	0.977
Automobile	0.967	0.969	0.973	0.987
Bird	0.887	0.917	0.937	0.960
Cat	0.822	0.856	0.898	0.931
Deer	0.912	0.937	0.946	0.972

Dog	0.857	0.876	0.914	0.938
Frog	0.936	0.950	0.969	0.981
Horse	0.930	0.947	0.961	0.978
Ship	0.961	0.957	0.975	0.985
Truck	0.955	0.960	0.964	0.984

5. Discussion

The Advanced CNN demonstrates strong performance for a scratch-built model (91.4% accuracy) but shows inefficiency due to its large parameter count (108M). EfficientNet-B0 performs better with fewer parameters due to compound scaling. ResNet-18, with residual connections, improves to nearly 95% accuracy. VGG16 leads with 96.9% accuracy, showing consistent and balanced performance across all classes.

6. Conclusion

Rank	Model	Accuracy (%)	Best Qualities
1	VGG16	96.92	Consistent performance, high generalization
2	ResNet-18	94.98	Strong residual learning, stable training
3	EfficientNet-B0	93.03	Lightweight and efficient
4	Advanced CNN	91.41	Solid baseline, can improve with BN and pooling

In conclusion, while all models achieved high overall accuracy, there were some subtle weaknesses in their predictions. The models struggled the most to distinguish between **cats and dogs**, likely due to the high visual similarity between these classes. Misclassifications were also observed between **birds and airplanes** and **trucks and cars** in the training set; however, these did not significantly impact the final test performance