Project Report : Multi-Class Image Classification using Pretrained ResNet-18 & VGG16

Project Name: Multi-Class Image Classification using Pretrained ResNet-18 and VGG16 (CIFAR-10 subset)
Course: GIKI Advanced AI Bootcamp 2025 — Week 02: RNN, Reinforcement Learning, & Computer Vision
By: Iftikhar Ali
Deadline: Tuesday | 12/08/2025

Problem Statement

The objective of this project is to perform multi-class image classification using transfer learning. We fine-tune two pretrained convolutional neural network architectures — ResNet-18 and VGG16 on a small subset of the CIFAR-10 dataset containing three classes: airplane, automobile, and bird. The goal is to train efficiently, evaluate model performance using multiple metrics, and compare results.

Dataset

Name: CIFAR-10 (subset of 3 classes: airplane, automobile, bird)
Source: Built-in from torchvision.datasets.CIFAR10 (no manual download required)
Size after filtering:

Train samples: 15,000

Test samples: 3,000

Reason for choice:

Small, clean dataset ideal for fast experimentation.

Balanced class distribution.

Easy to load without external downloads.

Dataset Loading Process

Used torchvision.datasets.CIFAR10 with automatic download.

Applied transforms: resize to 224×224 (to match pretrained model input), convert to tensor, normalize with ImageNet mean/std.

Filtered dataset to only keep classes 0, 1, and 2 (airplane, automobile, bird).

Created PyTorch DataLoaders for training and testing.

Preprocessing

Resize: 224×224

Normalization: mean [0.485, 0.456, 0.406], std [0.229, 0.224, 0.225]

Batch size: 32

No additional augmentation to keep training simple and fast.

Models

ResNet-18

Pretrained on ImageNet.

Last fully connected layer replaced with 3-class output.

Optimizer: Adam, LR=0.001.

VGG16

Pretrained on ImageNet.

Last classifier layer replaced with 3-class output.

Optimizer: Adam, LR=0.001.

Training Procedure

Epochs: 10 for both models.

Loss function: CrossEntropyLoss.

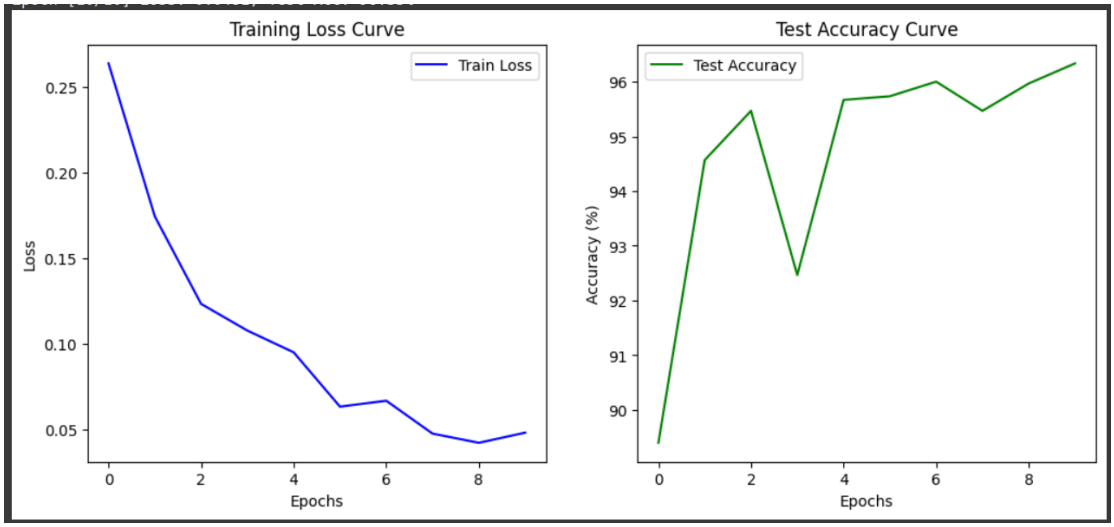
Trained on GPU.

Recorded training loss and test accuracy after each epoch.

Used a quick_fit_check function to compare train vs. test accuracy for overfitting/underfitting detection.

-

Results
Training Summary
ResNet-18
Train Accuracy: **99.33%**
Test Accuracy: **96.33%**
Verdict: Balanced (no strong signs of over/underfitting).

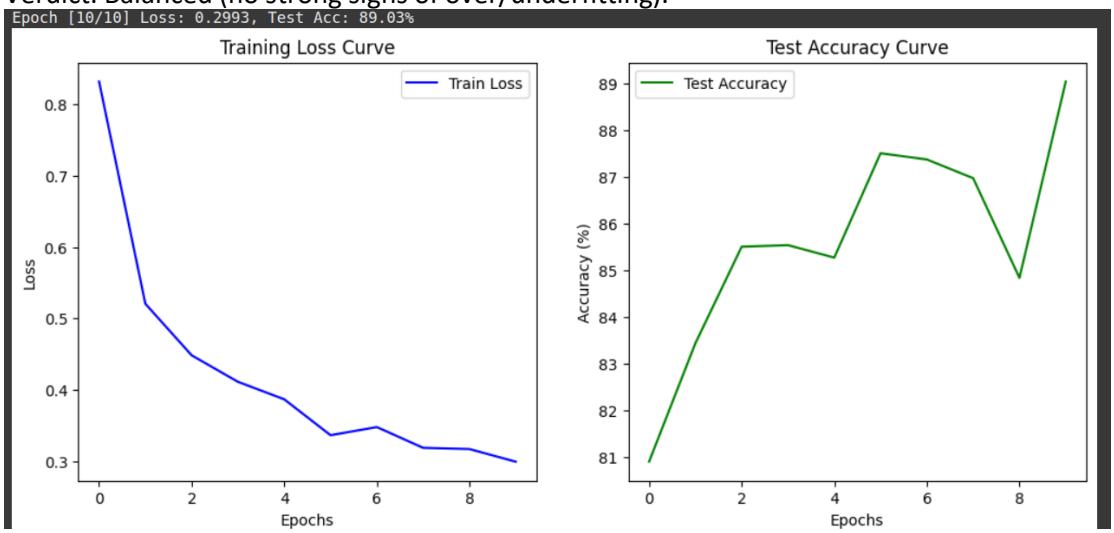


VGG16

Train Accuracy: **91.32%**

Test Accuracy: **89.03%**

Verdict: Balanced (no strong signs of over/underfitting).



Evaluation Metrics

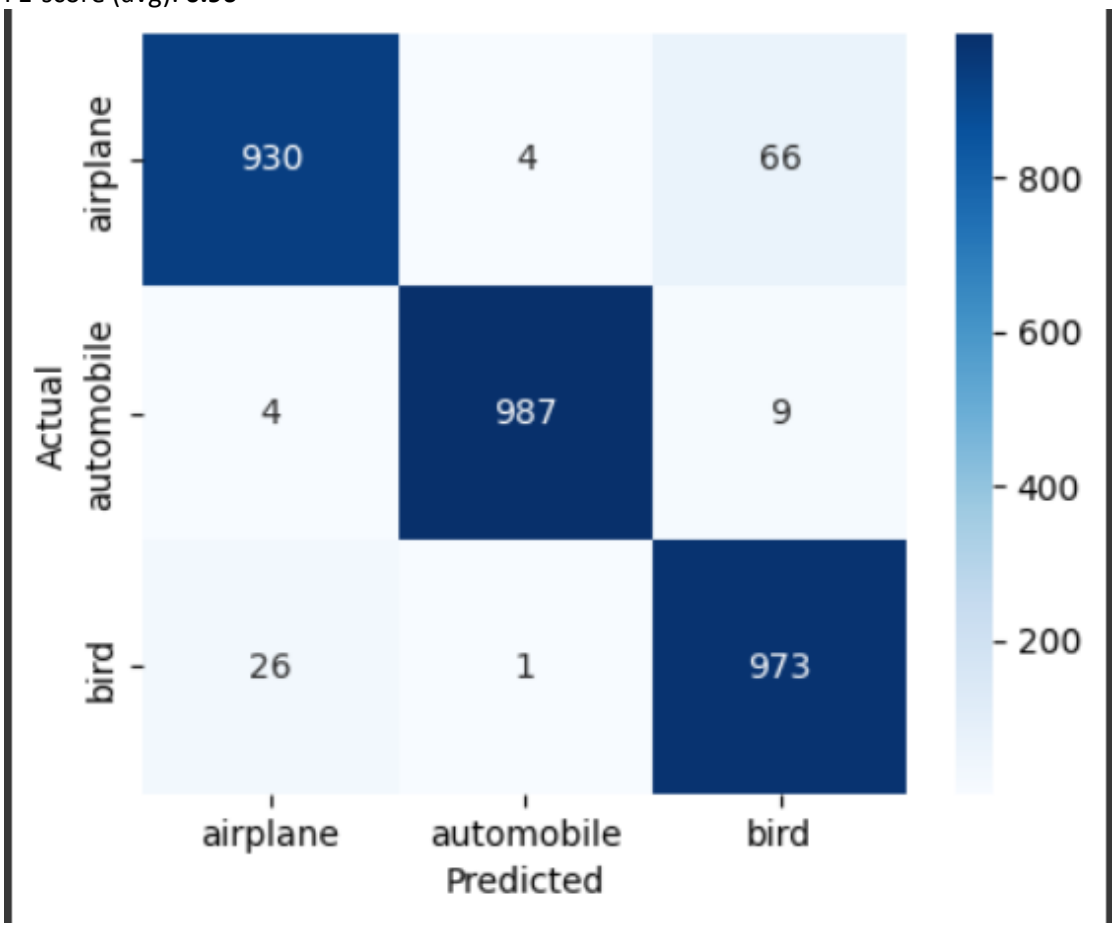
ResNet-18

Test Accuracy: **96.33%**

Precision (avg): **0.96**

Recall (avg): **0.96**

F1-score (avg): **0.96**



Classification Report (ResNet-18):

Category	Precision	Recall	F1-Score
airplane	0.97	0.93	0.95
automobile	0.99	0.99	0.99
bird	0.93	0.97	0.95

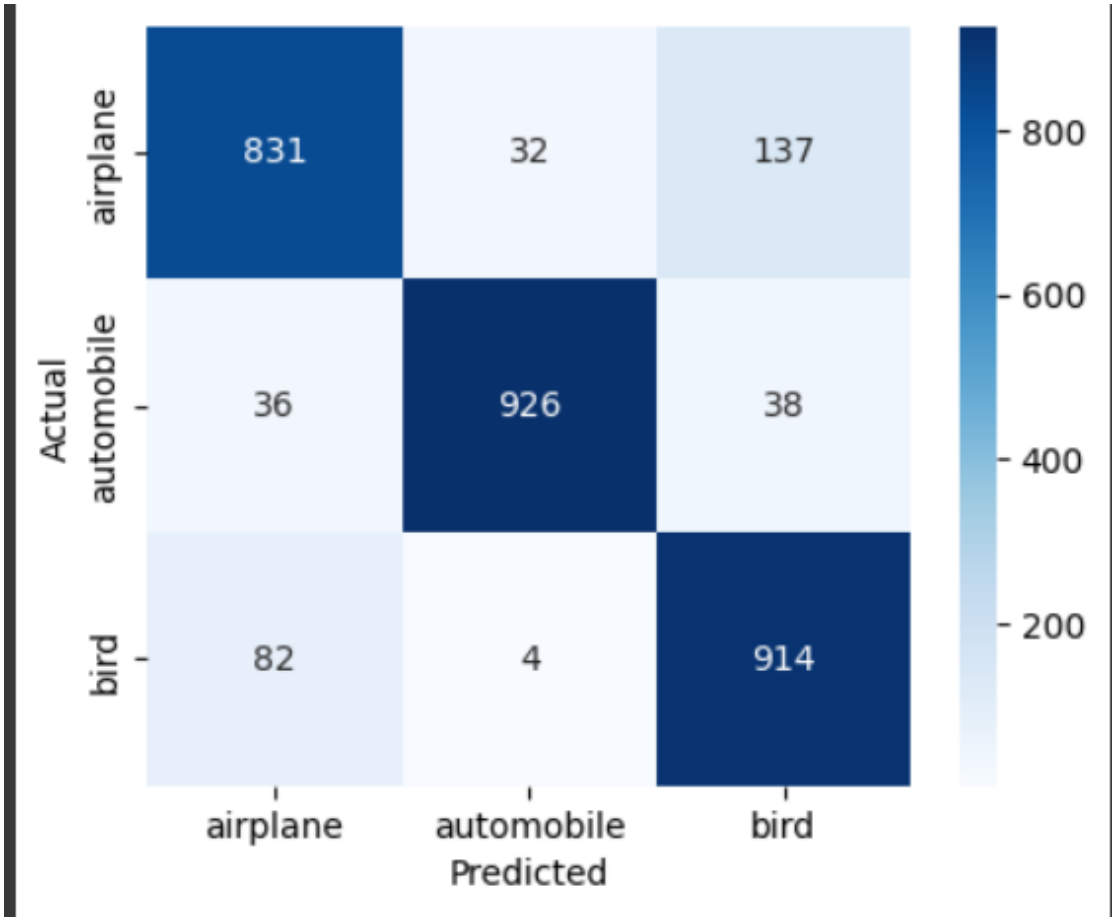
VGG16

Test Accuracy: **89.03%**

Precision (avg): **0.89**

Recall (avg): **0.89**

F1-score (avg): **0.89**



Classification Report (VGG16):

Category	Precision	Recall	F1-Score
airplane	0.88	0.83	0.85
automobile	0.96	0.93	0.94
bird	0.84	0.91	0.88

Comparison Table

Model	Train Acc	Test Acc	Avg Precision	Avg Recall	Avg F1
ResNet-18	99.33%	96.33%	0.96	0.96	0.96
VGG16	91.32%	89.03%	0.89	0.89	0.89

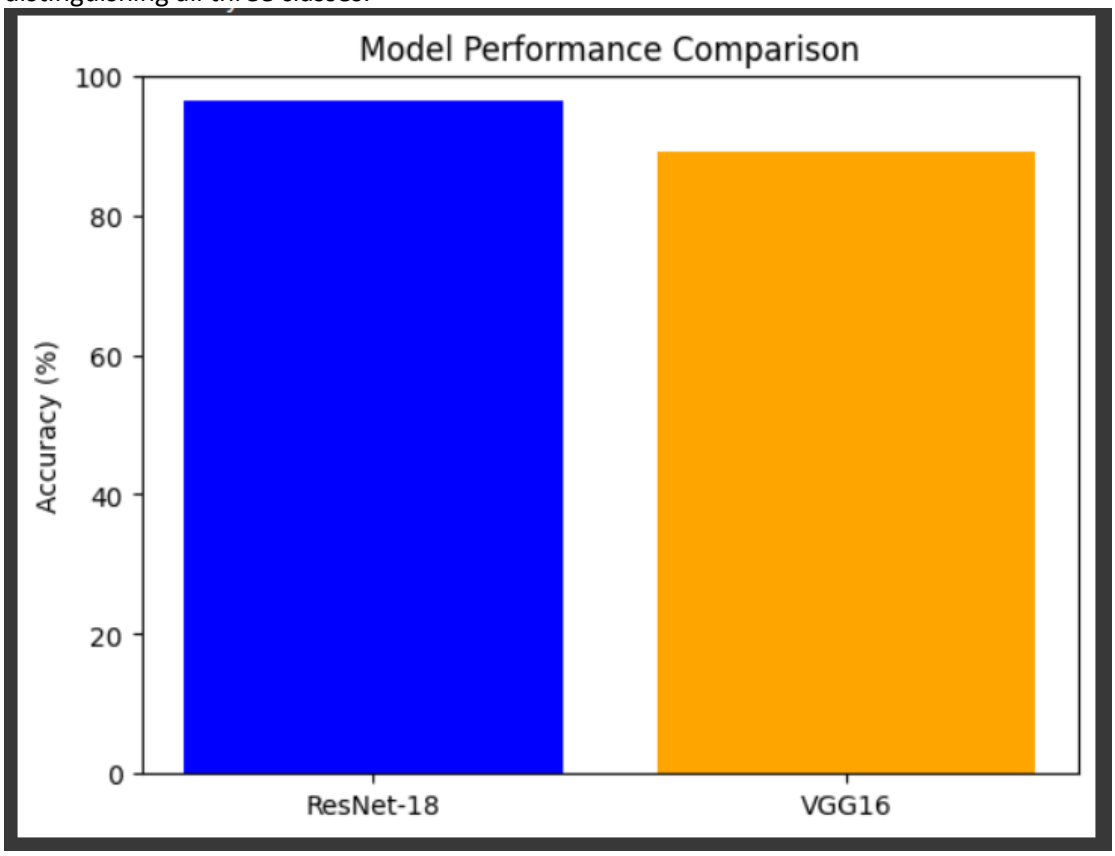
Insights

ResNet-18 outperformed VGG16 in both accuracy and all classification metrics.

ResNet-18 trained faster and generalized better on this small dataset.

No major overfitting signs were observed for either model — train and test accuracy were close.

Confusion matrices showed minimal class confusion, with ResNet-18 especially strong at distinguishing all three classes.



-

Conclusion

We successfully fine-tuned ResNet-18 and VGG16 for a 3-class subset of CIFAR-10.
ResNet-18 achieved **96.33%** test accuracy and outperformed VGG16 in all metrics.
For this dataset and hardware constraints, ResNet-18 is the recommended choice.