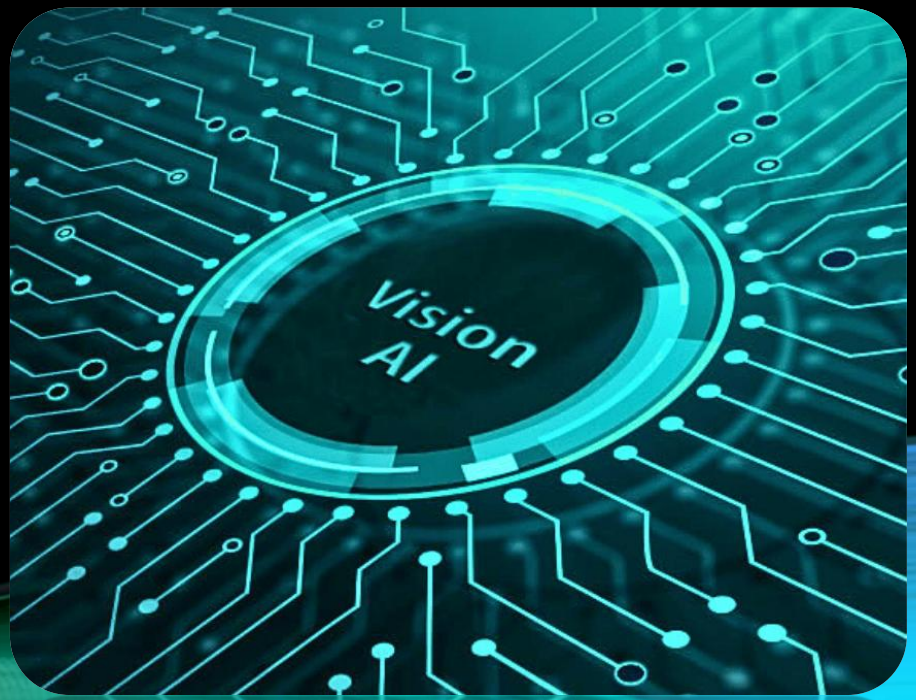
VISION AI – CIFAR-10 IMAGE CLASSIFICATION

Using CNN & MobileNetV2 Transfer Learning

By, Dipayan Majumdar



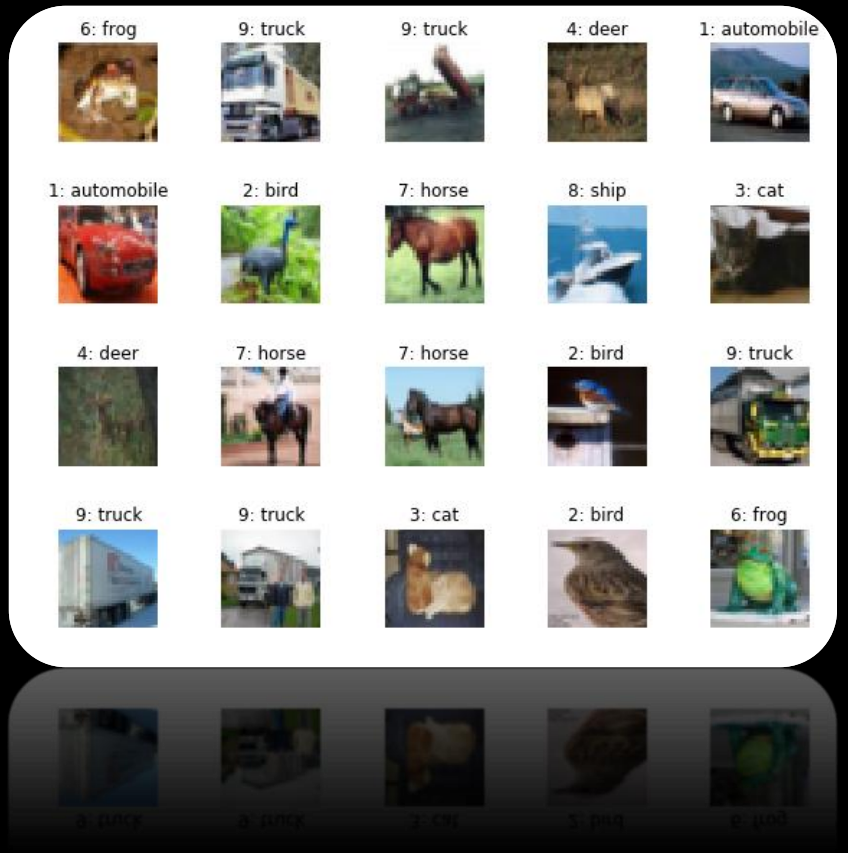
Highlights

- **Dataset:** CIFAR-10 (60,000 images, 10 categories)
- **Tech Stack:** Python, TensorFlow, Keras, Google Colab
- **Key Skills:** Data preprocessing, Deep Learning, Transfer Learning, Model Evaluation

DATASET & PREPROCESSING

Dataset: CIFAR-10

- **Size:** 60,000 images (32×32 pixels)
- **Classes:** Airplane, Automobile, Bird, Cat, Deer, Dog, Frog, Horse, Ship, Truck
- **Train:** 45,000 | **Validation:** 5,000 | **Test:** 10,000
- Source: CIFAR-10 Dataset



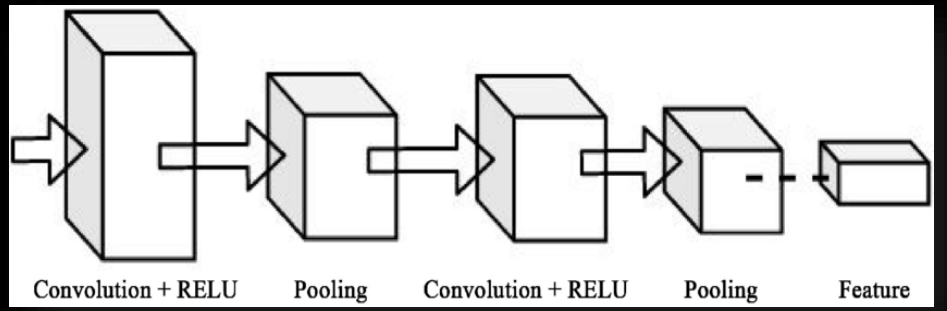
Preprocessing:

- Normalized pixel values to [0,1]
- Train/validation/test split
- Applied augmentation: rotation, width/height shift, horizontal flip

MODELS & METHOD

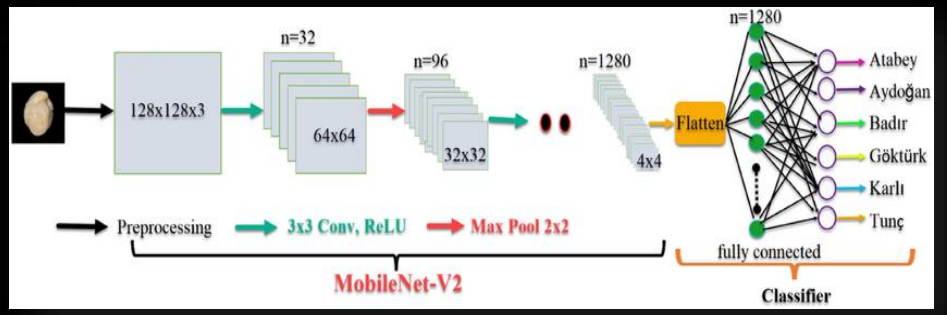
1. Baseline CNN

- 3 Conv2D layers + MaxPooling
- Dense layers for classification
- Adam optimizer, 10 epochs



2. MobileNetV2 Transfer Learning

- Pretrained on ImageNet
- On-the-fly resizing to 160×160
- Frozen base layers + trained dense head
- Fine-tuned last 30 layers
- Adam optimizer (lr=1e-5), 5+5 epochs

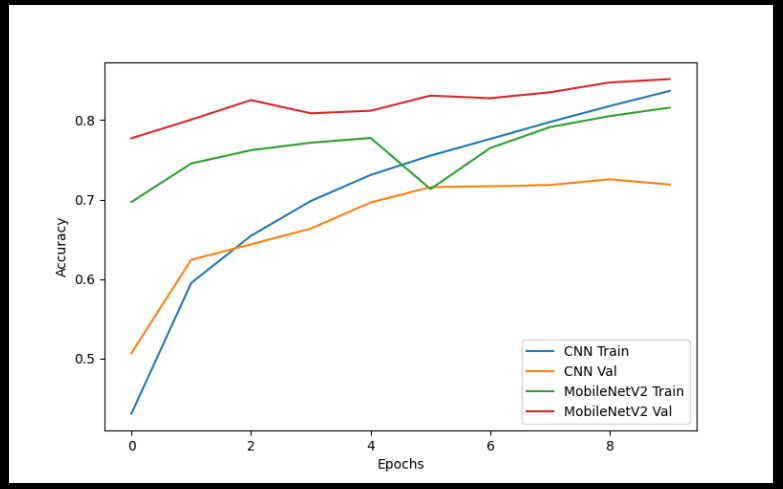


RESULTS

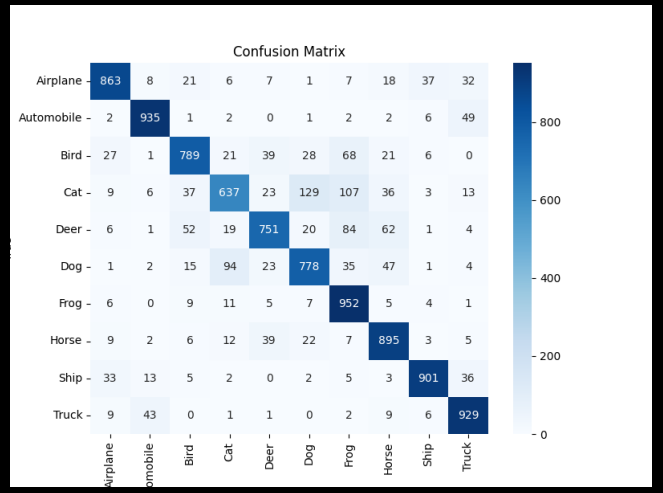
Accuracy:

Model	Val Accuracy	Test Accuracy
Baseline CNN	~72%	~72%
MobileNetV2 (fine-tuned)	~85%	~84%

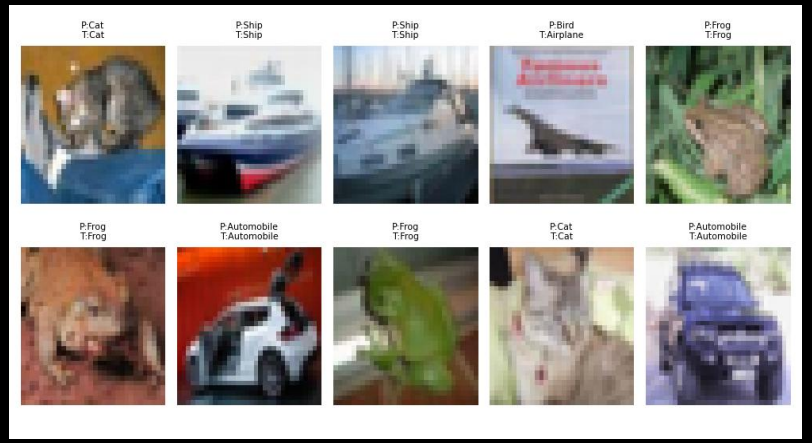
Visuals:



Accuracy Curves



Sample Predictions



Confusion Matrix

CONCLUSION

Key Takeaways

- MobileNetV2 with fine-tuning achieved **84% test accuracy**
- Significant improvement over baseline CNN
- Demonstrated skills in preprocessing, CNN architecture, transfer learning, and evaluation

Future Improvements

- Train longer for higher accuracy
- Try EfficientNet or ResNet architectures
- Experiment with advanced augmentation