DEEP LEARNING:

Image Classification using CNN and Transfer Learning

Chantal Silva
Diogo Simões
Javier Garcia

14 November 2025

1. Dataset & Objective

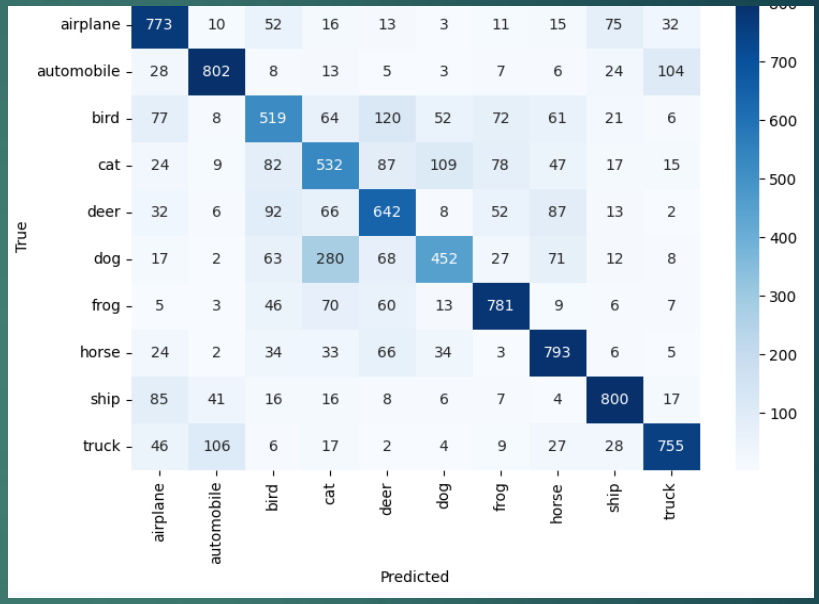
- • CIFAR-10: 60,000 color images (32×32 px)
- • 10 classes: airplane, car, bird, cat, deer, dog, frog, horse, ship, truck
- • Split: 50,000 train / 10,000 test
- • Goal: build and compare models to classify images correctly

2. Project Workflow

- 1. Data loading & preprocessing
- 2. Train baseline CNN from scratch
- 3. Design improved CNN+ architecture
- 4. Apply transfer learning (MobileNetV2 & NASNetMobile)
- 5. Evaluate with accuracy, precision, recall, F1
- 6. Analyze and compare models

3. Baseline CNN (from scratch)

- **Architecture:**
- • Input (32×32×3)
- • 2× Conv2D + MaxPooling blocks
- • Flatten → Dense(64) + Dropout(0.3)
- • Output Dense(10, softmax)
- **Metrics (test set):**
  - • Accuracy: 0.685
  - • Precision (macro): 0.686
  - • Recall (macro): 0.685
  - • F1-score (macro): 0.683
  - • Confusion matrix: cats vs. dogs and some animals are often confused



	Precision	Recall	F1
airplane	0.7	0.77	0.73
automobile	0.81	0.8	0.81
bird	0.57	0.52	0.54
cat	0.48	0.53	0.5
deer	0.6	0.64	0.62
dog	0.66	0.45	0.54
frog	0.75	0.78	0.76
horse	0.71	0.79	0.75
ship	0.8	0.8	0.8
truck	0.79	0.76	0.77

4. CNN + (BN + GAP + Regularization)

- **Key improvements over baseline:**
- • Data augmentation: flips, rotations, zoom, contrast
- • Batch Normalization after conv layers
- • Global Average Pooling instead of Flatten
- • L2 regularization + Dropout(0.5)
- **Metrics (test set):**
  - • Accuracy: 0.871
  - • Precision (macro): 0.870
  - • Recall (macro): 0.871
  - • F1-score (macro): 0.870
  - • Confusion matrix: large reduction in animal vs. animal confusion

	Precision	Recall	F1
airplane	0.88	0.88	0.88
automobile	0.93	0.95	0.94
bird	0.82	0.86	0.84
cat	0.8	0.72	0.76
deer	0.85	0.85	0.85
dog	0.85	0.77	0.81
frog	0.87	0.94	0.9
horse	0.89	0.9	0.89
ship	0.93	0.91	0.92
truck	0.88	0.94	0.91

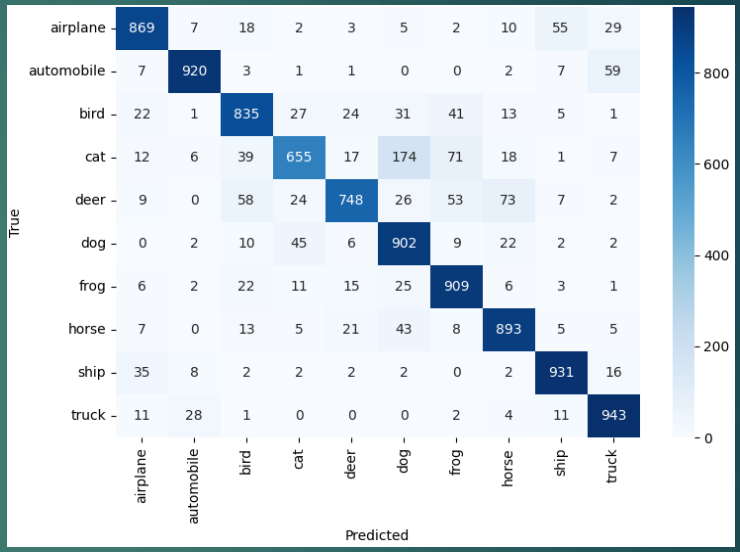
5. MobileNetV2 (Transfer Learning)

- **Pipeline:**
- • Resize images to 160×160
- • Use pretrained MobileNetV2 (ImageNet) as frozen base
- • Add new classification head (GAP → Dropout → Dense(10))
- • Warm-up: train only the head
- • Fine-tune: unfreeze top ~30 layers with lower learning rate
- **Metrics (test set):**
  - • Accuracy: 0.897 (best model)
  - • Precision (macro): 0.898
  - • Recall (macro): 0.897
  - • F1-score (macro): 0.896
  - • Confusion matrix: clearer diagonal, especially for vehicles and ships

	Precision	Recall	F1
airplane	0.91	0.9	0.9
automobile	0.94	0.94	0.94
bird	0.94	0.84	0.89
cat	0.81	0.78	0.8
deer	0.9	0.89	0.9
dog	0.88	0.82	0.85
frog	0.84	0.97	0.9
horse	0.94	0.91	0.92
ship	0.93	0.94	0.94
truck	0.89	0.96	0.92

6. NASNetMobile (Transfer Learning)

- **Pipeline:**
- • Resize images to 224×224
- • Pretrained NASNetMobile as base model
- • New classification head on top
- • Warm-up: train head with base frozen
- • Fine-tune: unfreeze last ~20 layers with small learning rate
- **Metrics (test set):**
  - • Accuracy: 0.861
  - • Precision (macro): 0.863
  - • Recall (macro): 0.861
  - • F1-score (macro): 0.859
  - • Confusion matrix: similar to MobileNetV2, slightly lower for some classes

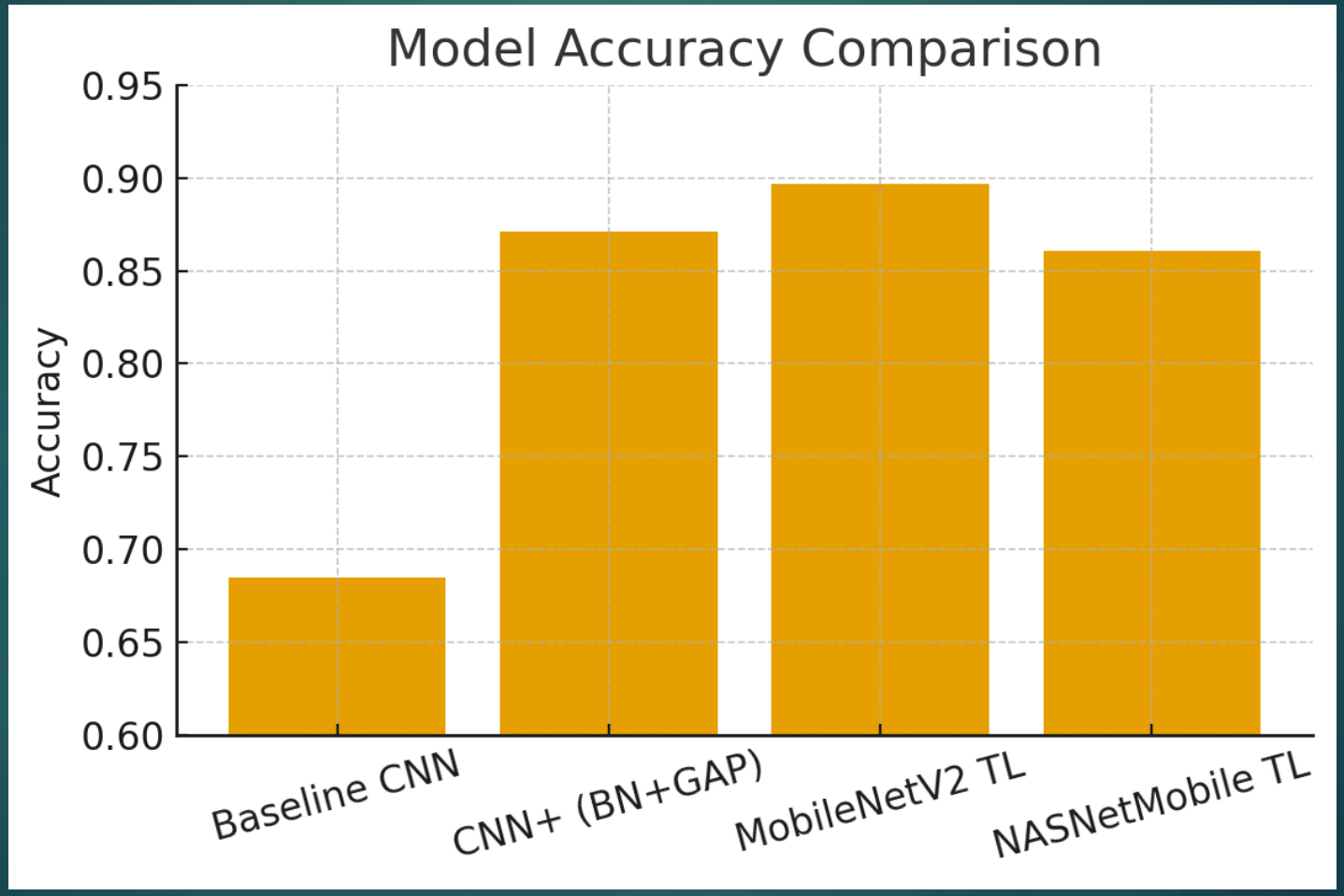


	Precision	Recall	F1
airplane	0.89	0.87	0.88
automobile	0.94	0.92	0.93
bird	0.83	0.83	0.83
cat	0.85	0.66	0.74
deer	0.89	0.75	0.81
dog	0.75	0.9	0.82
frog	0.83	0.91	0.87
horse	0.86	0.89	0.87
ship	0.91	0.93	0.92
truck	0.89	0.94	0.91

7. Global Metrics per Model (Macro Averages)

Model	Accuracy	Precision	Recall	F1-score
Baseline CNN	0.6849	0.6859	0.6849	0.6826
CNN+ (BN+GAP)	0.8710	0.8703	0.8710	0.8699
MobileNetV2 TL	0.8968	0.8979	0.8968	0.8963
NASNetMobile TL	0.8605	0.8634	0.8605	0.8590

8. Visual Comparison of Accuracy

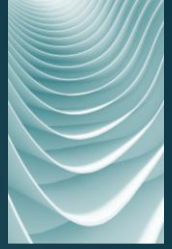


9. Final Comparison & Takeaways

- Baseline CNN provides a reference point (~68% accuracy).
- CNN+ shows the impact of better architecture & regularization (~87%).
- Transfer learning models achieve the best results (MobileNetV2 ≈ 90%).
- Confusion matrices show that transfer learning reduces class confusion, especially between visually similar animal classes.
- Trade-off: transfer learning is more accurate but also more computationally expensive.

10. Conclusion

- Demonstrated full deep learning workflow on CIFAR-10.
- Compared custom CNNs with modern transfer learning models.
- Best model: MobileNetV2 (transfer learning, fine-tuned).
- Educational takeaway: start with a simple CNN, then move to more advanced architectures and transfer learning for higher performance.



Thank you!