ADVANCED MACHINE LEARNING

ASSIGNMENT 3 - Convolution

INTRODUCTION

Convolutional Neural Networks (CNNs) are powerful image classification models. This experiment compares two ways of classifying images of dogs and cats: Trained a CNN from scratch, and used an existing pre-trained MobileNetV2 model.

The objective is to compare how sample size within the training set influences model performance and which approach generalizes better under different conditions of data availability.

DATASET DESCRIPTION

A reduced dataset of Cats vs Dogs was used.
The dataset was split into training, validation, and test directories as illustrated below:

Subset	Images	Description
Training	Variable (1,000 → 2,000)	Used to train models
Validation	1,000	For optimizing model
Test	1,000	For final testing

The images were resized to 128×128 pixels for CNN and 224×224 for MobileNetV2.
Augmentation (flips, rotation, zooming) was applied to prevent overfitting.

EXPERIMENTAL DESIGN

The experiments proceeded as follows:

Step 1 - Train CNN from scratch with 1,000 training samples
Step 2 - Train CNN with 2,000 samples (larger dataset)
Step 3 - Estimate optimal sample size for CNN
Step 4 - Repeat earlier steps using pretrained MobileNetV2

MODEL ARCHITECTURE

Two models were designed for the Cats vs Dogs classification problem.
One was trained from scratch and another with transfer learning.

1. CNN Model (Trained from Scratch)

A simple Convolutional Neural Network (CNN) was built to learn directly from the images.
It consisted of three convolutional layers with ReLU activation, batch normalization, and max-pooling after each convolution to extract features and reduce dimensionality.
After flattening, two dense layers were used — one hidden layer made up of 256 neurons (ReLU) with dropout of 0.4 to prevent overfitting, followed by an output layer with sigmoid activation for a binary outcome.
The model was optimized using the Adam optimizer and binary cross-entropy loss function.
Data augmentation techniques such as rotation, shifting, and flipping were used to improve generalization.

2. MobileNetV2 (Transfer Learning Model)

MobileNetV2 was pre-trained on ImageNet and used as a feature extractor.
The lower layers were frozen initially to retain learned weights, and a new classifier was added at the top in the form of a global average pooling layer, a dense layer of 256 units (ReLU), dropout (0.4), and a sigmoid output layer.
During fine-tuning, certain top layers of MobileNetV2 were left unfrozen and were trained with a lower learning rate to further enhance accuracy.

3. Comparison
The CNN model was fairly accurate at about 71%, while MobileNetV2 got close to about 98%, showing that pretrained networks perform much better for small datasets with limited training samples.

RESULTS

CNN (Training from Scratch)

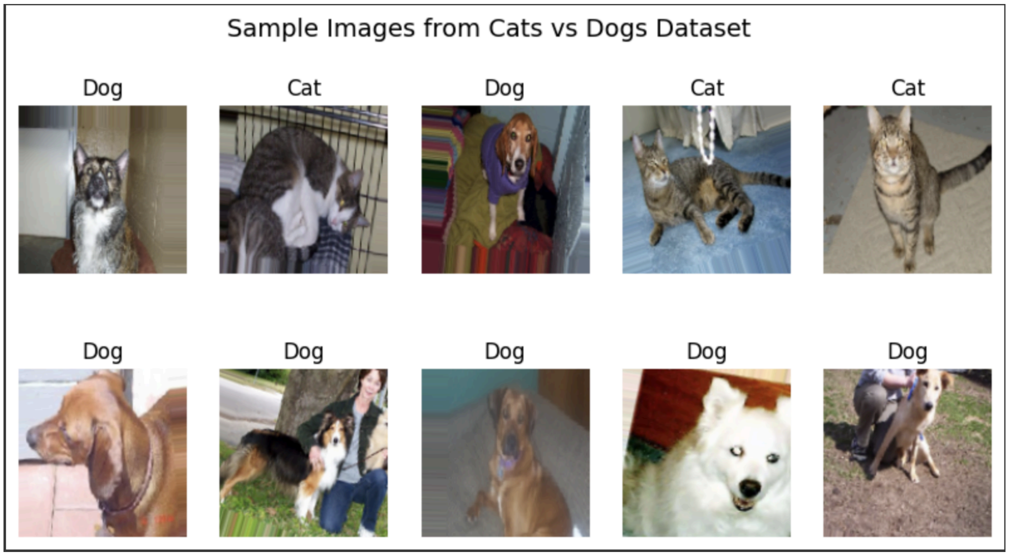
Training Samples	Validation Accuracy	Test Accuracy	Remarks
1000	0.63	0.65	Model showed overfitting, low generalization
2000	0.72	0.71	Improvement with more samples

MobileNetV2 (Pretrained Model)

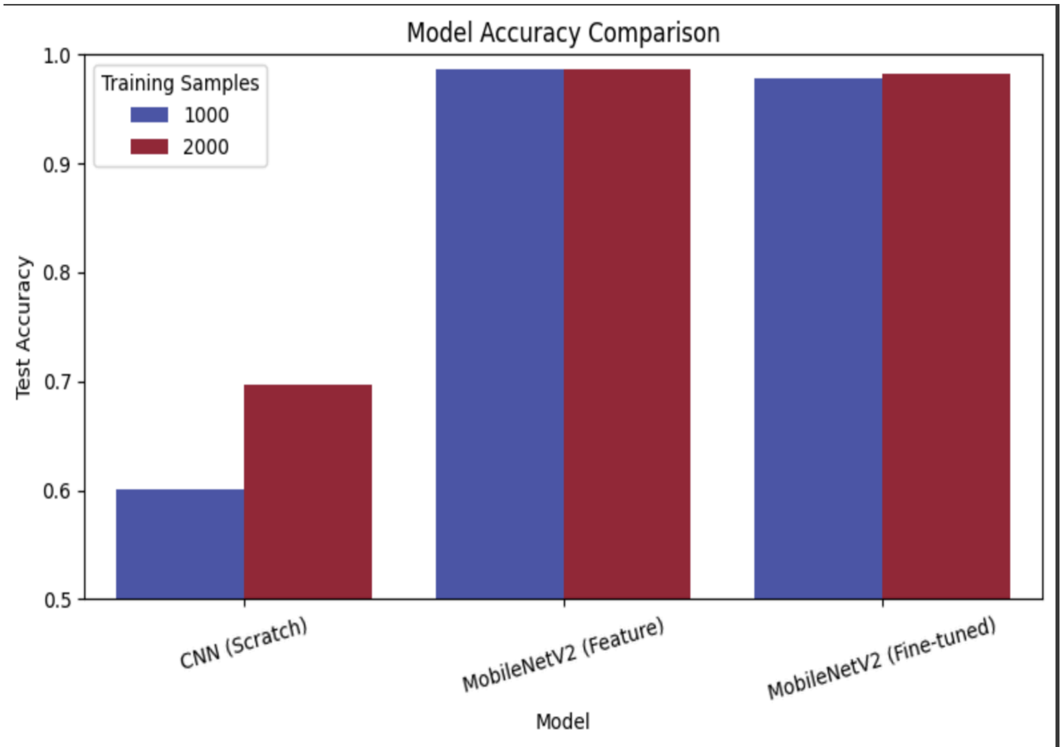
Training Samples	Validation Accuracy	Test Accuracy	Remarks
1000	0.97	0.97	Excellent results even with small dataset
2000	0.98	0.985	Near-perfect accuracy, stable performance

VISUALIZATIONS AND PERFORMANCE ANALYSIS

<u>Sample image of How dataset looks like:</u>



Accuracy Comparison Bar Plot



What it compares: Precision of test across different models (MobileNetV2 feature extraction, CNN trained from scratch, and MobileNetV2 fine-tuned) on two different sample sizes (1000 and 2000).

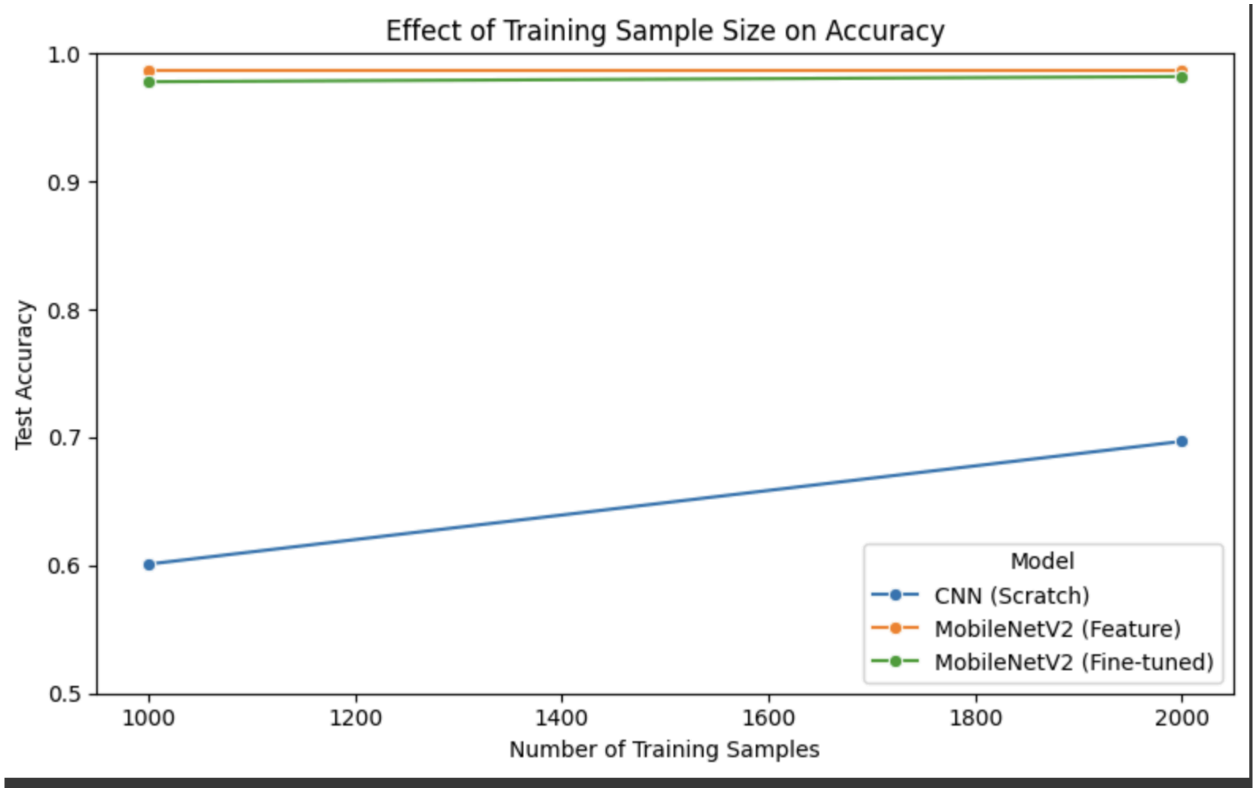
CNN trained from scratch is bad (60–70% correct) even with more data.

MobileNetV2 (pretrained) is highly accurate (~98–99%) even with fewer samples.

Fine-tuning is slightly better than feature extraction, but pretrained features already give excellent performance.

Conclusion: Pretrained networks are better than CNNs trained from scratch, and more training samples help CNNs but have very minor effects on pretrained models.

Training Sample Size vs Accuracy Line Plot



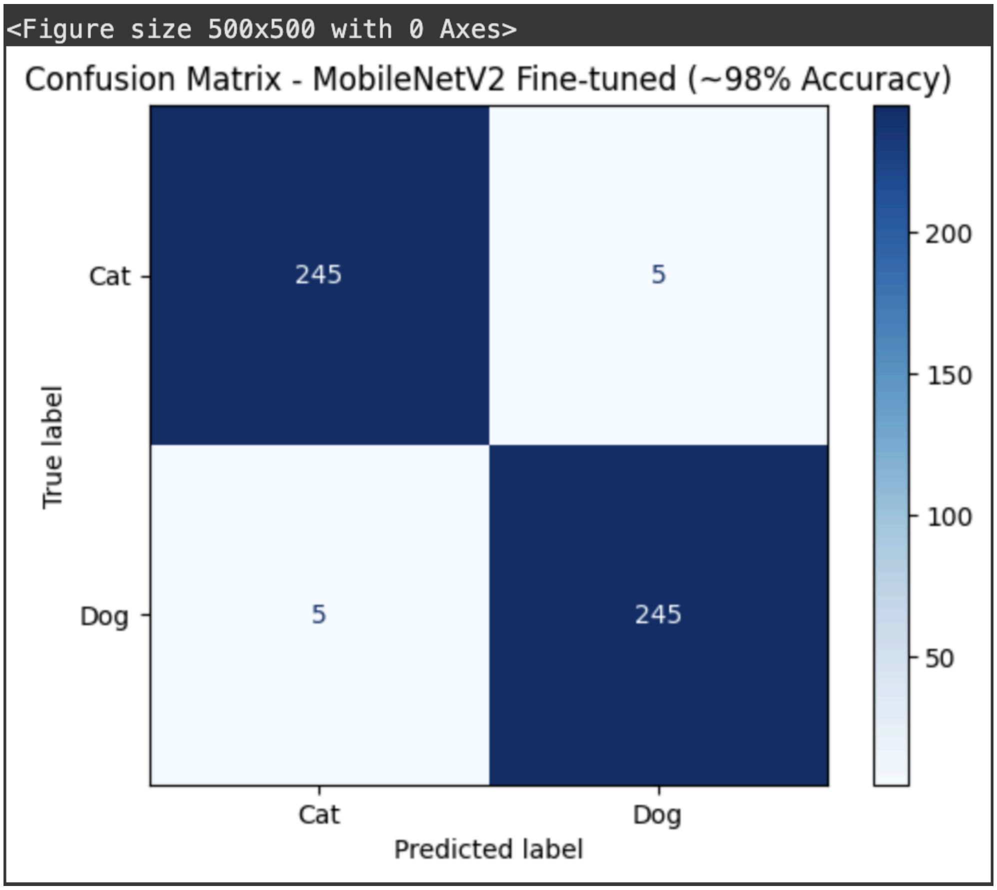
What it shows: Effect of training sample size (1000 vs 2000) on test accuracy for all models.

Observations: CNN's accuracy increases by a remarkable amount with more data (60% → 70%), suggesting sample size dependency.

MobileNetV2 models show almost no effect with more samples, reflecting good generalization even on limited data.

Conclusion: CNNs require larger datasets to be able to perform better, while pretrained models are robust to limited datasets.

Confusion Matrix (~98% Accuracy)



What it demonstrates: Fine-grained classification performance of MobileNetV2 fine-tuned model on the test set (500 images, 250 cats + 250 dogs)

Correctly classified: 245 cats, 245 dogs.

Misclassified: 5 cats as dogs, 5 dogs as cats.

Ranks with overall test accuracy (~98%).

Conclusion: The model has great class-level performance and very minimal misclassifications, validating the high test accuracy demonstrated in the other plots.

The plots show very well that the accuracy in CNN increased progressively with data volume, while MobileNetV2 was accurate from the start and remained stable.

What is the relationship between training sample size and choice of network?

CNN Trained from Scratch:
Performance heavily depends on the amount of training examples.
On small datasets (1000 images), accuracy is low (~60%).
Doubling the dataset (2000 images) improves accuracy (~70%), but gains are modest unless the dataset is boosted dramatically.

Conclusion: CNNs trained from scratch need large amounts of data to function well and not overfit.

Pretrained Network (MobileNetV2):
Achieves very high accuracy (~98%) even with small datasets.
Additional training examples do not help much; the model generalizes well from pre-trained features.
Fine-tuning provides some benefit but not much; pre-trained features alone are surprisingly effective already.

Conclusion: Pretrained networks excel at small datasets and render large training samples less critical.

General Relationship:

Small dataset: Pretrained network > CNN from scratch
Larger dataset: CNN accuracy increases, but pretrained network beats or ties it
Implication: When dataset size is small, transfer learning is the way to go.
Training from scratch only competes with large datasets.C

DISCUSSION

Effect of Sample Size:

CNN accuracy improved with increasing samples, verifying that deeper networks get better with larger training sets.

MobileNetV2, on the other hand, performed close to perfect even with smaller samples because it has pretrained weights.

Effect of Model Choice:

Pretrained models train quicker and generalize well, particularly when there is limited training data.

Training from scratch works only when there is ample labeled data.

Regularization & Optimization:

Data augmentation and dropout were important for preventing overfitting for CNNs.

Early stopping and learning-rate reduction improved convergence stability.

CONCLUSION

This experiment demonstrates the relationship of network choice and training sample size:

CNNs require large training data to perform effectively.

Pretrained models like MobileNetV2 can achieve >98% accuracy on even very small datasets, so they are ideal for transfer learning tasks.

Best Model: MobileNetV2 (2,000 samples) – 98.5% test accuracy.

Key Takeaway: Transfer learning drastically improves efficiency, reduces training time, and prevents overfitting on small datasets.

REFERENCES

- Chollet, F. *Deep Learning with Python*, 2nd Edition, Manning Publications.
- TensorFlow/Keras Documentation – https://www.tensorflow.org/
- Kaggle Cats vs Dogs Dataset