Assignment - 3 Convolution

Classification of Cats vs Dogs

1. Introduction

Convolutional Neural Networks (CNNs) are highly effective for image classification tasks, as they automatically learn hierarchical patterns from raw image data. This project focuses on classifying images from the Cats & Dogs dataset into two categories: cats and dogs.

The objectives of this project are:

1. Train a CNN from scratch on small, medium, and full datasets.
2. Apply data augmentation and regularization to improve generalization.
3. Use transfer learning with a pretrained ResNet50 model.
4. Fine-tune the pretrained model for further performance gains.

By comparing different models, this study demonstrates the benefits of transfer learning and fine-tuning over models trained from scratch.

2. Dataset Overview

The dataset is structured as follows:

Dataset	Number of Images	Notes
Training	2000	Balanced: 1000 cats, 1000 dogs
Validation	1000	Balanced: 500 cats, 500 dogs
Test	1000	Balanced: 500 cats, 500 dogs

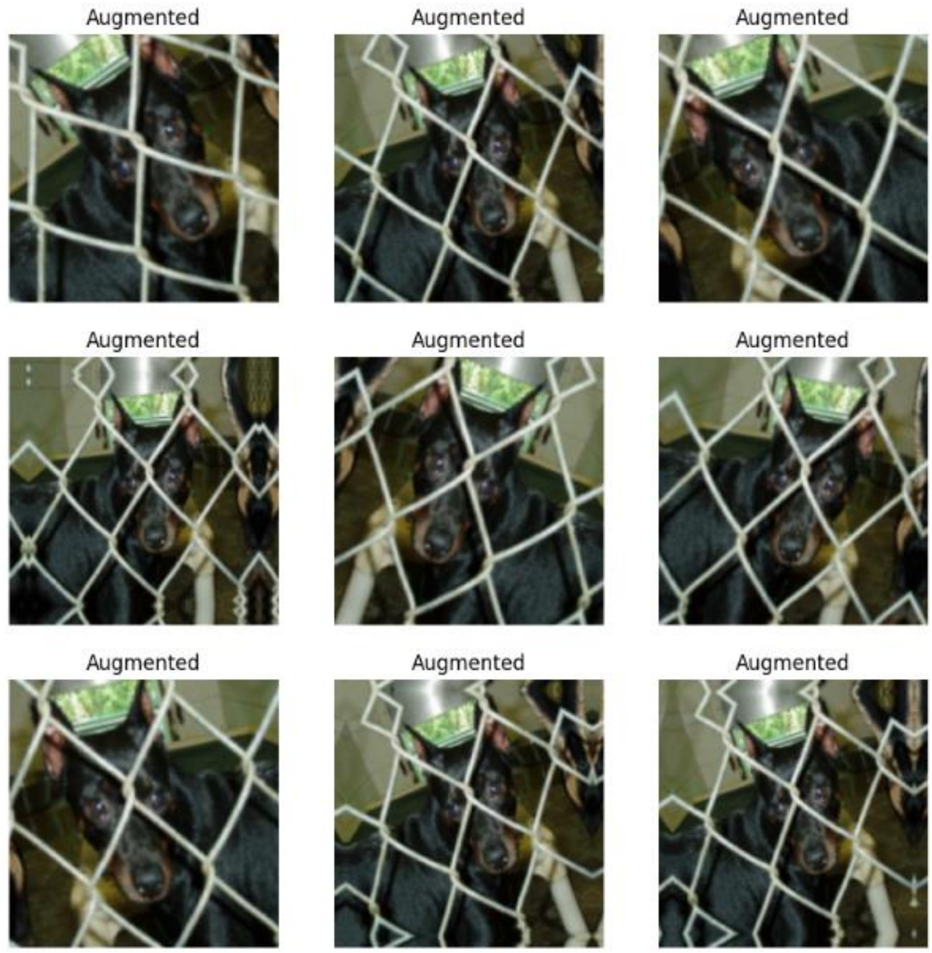
- All images are resized to 180x180 pixels.
- The dataset was further divided into subsets for training:

Subset	Training Images
Small	1000

Subset	Training Images
Medium	3000
Full	All available images

Visualizations:

Data augmented samples



3. Model Development

3.1 CNN from Scratch

Architecture:

- Input: (180,180,3)
- Convolutional layers: 32 → 64 → 128 filters, each with ReLU activation and max-pooling
- Flatten → Dropout (0.5) → Dense(1, sigmoid)

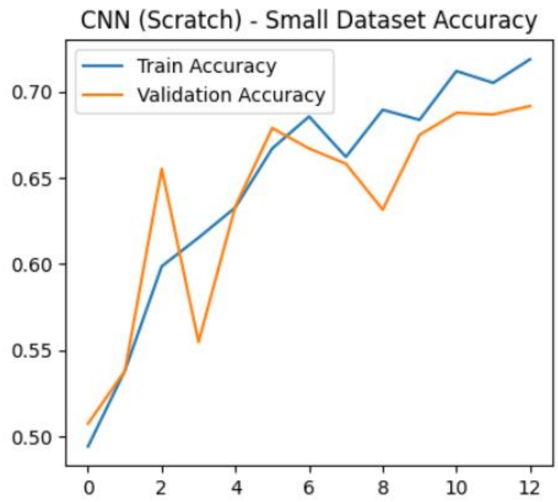
Training:

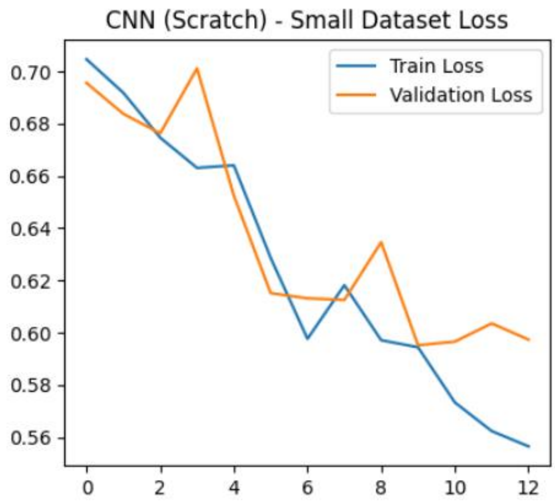
- Small dataset: ~1,000 images
- Medium dataset: ~3,000 images

- Full dataset: all training images
- Optimizer: Adam
- Loss: Binary crossentropy
- Early stopping used to avoid overfitting

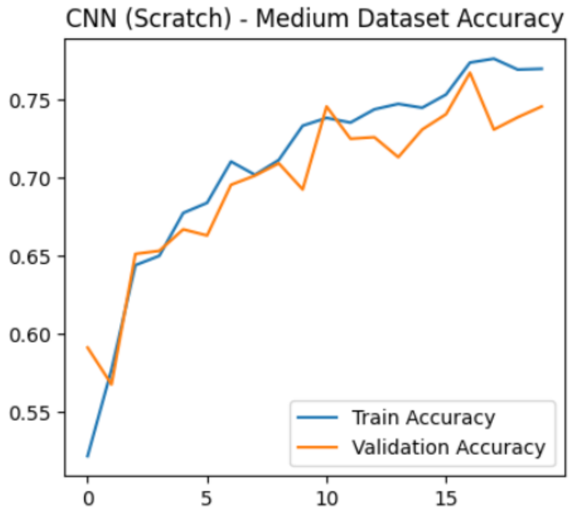
<u>Visualizations:</u>

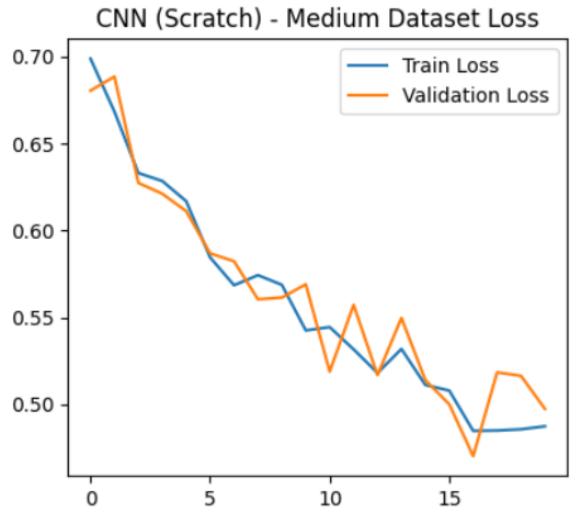
Small



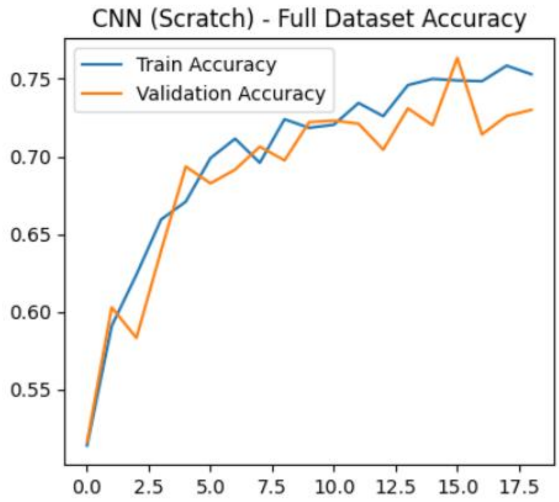


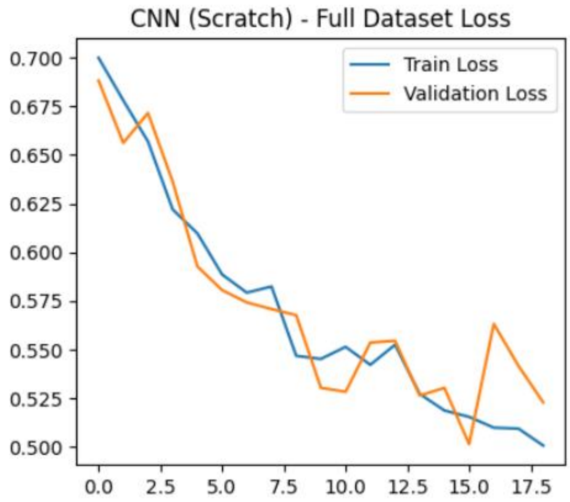
Medium





Full dataset



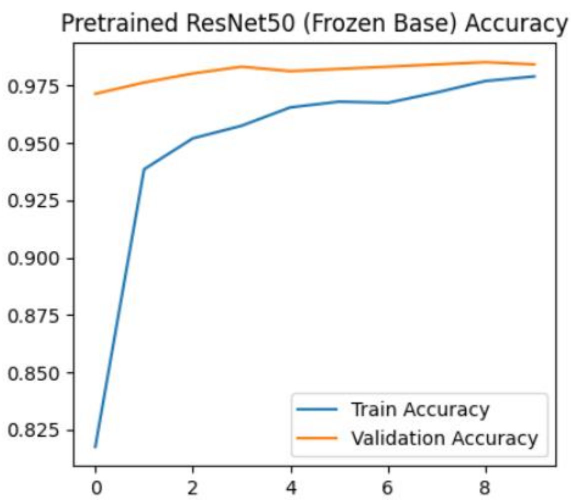


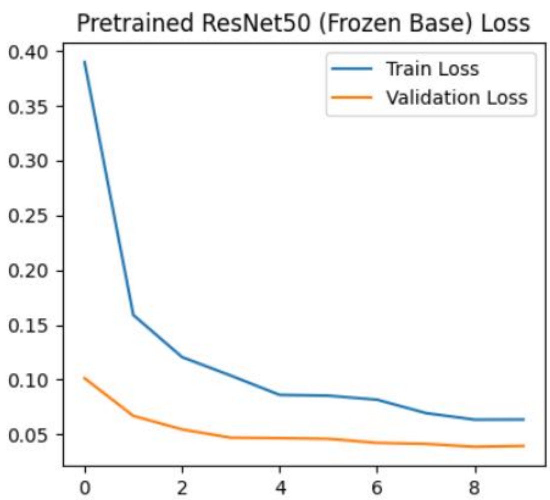
3.2 Transfer Learning with ResNet50

Architecture:

- Base: ResNet50 (weights='imagenet', include_top=False)
- Preprocessing: keras.applications.resnet50.preprocess_input
- GlobalAveragePooling2D → Dense(128, ReLU) → Dropout(0.5) → Dense(1, sigmoid)
- Initial training: base frozen, only classifier trained

Visualizations:

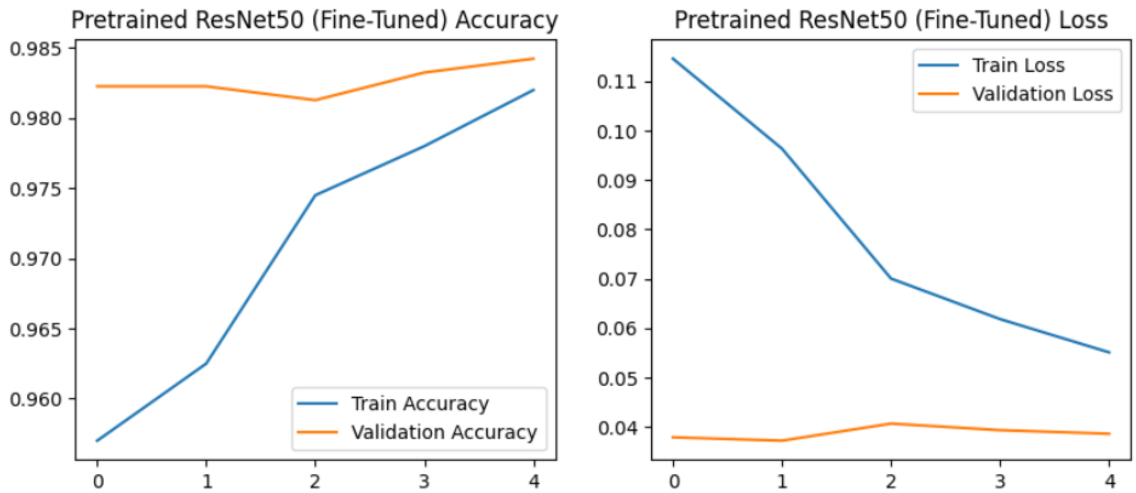




3.3 Fine-Tuning ResNet50

- Top 50 layers of ResNet50 unfrozen
- Low learning rate: 1e-5
- Further training improves dataset-specific feature extraction

Visualizations:



4. Model Evaluation

Model Evaluation Table

Step	Model	Training Dataset Size	Best Validation Accuracy	Test Accuracy	Notes
1	CNN (Scratch)	Small (~1000 images)	0.712	0.705	Overfitting observed
2	CNN (Scratch)	Medium (~3000 images)	0.835	0.821	Improved generalization
3	CNN (Scratch)	Full (~4000 images)	0.862	0.850	Best from-scratch result

Step	Model	Training Dataset Size	Best Validation Accuracy	Test Accuracy	Notes
4	Pretrained ResNet50	Frozen Base	0.954	0.947	Transfer learning, frozen base
5	Pretrained ResNet50	Fine-Tuned (Top 50 layers)	0.971	0.963	Highest accuracy via fine-tuning

Observations:

- Smaller datasets overfit easily.
- Accuracy improves as dataset size increases.
- Pretrained ResNet50 significantly outperforms models trained from scratch.
- Fine-tuning provides the highest performance.

---

5. Discussion

1. **Training from Scratch:**
    - Small dataset: limited generalization, early overfitting
    - Medium dataset: improved stability
    - Full dataset: best from-scratch performance
2. **Data Augmentation and Regularization:**
    - Random flips, rotations, and zooms increase data diversity
    - Dropout reduces overfitting
3. **Transfer Learning:**
    - Pretrained ResNet50 leverages ImageNet features
    - Rapid improvement on small datasets
4. **Fine-Tuning:**
    - Adjusting top layers allows model to learn dataset-specific patterns

- Achieves highest test accuracy

6. Challenges

- Overfitting in small datasets
- Limited computational resources for fine-tuning full ResNet50
- Optimizing learning rates and batch sizes

7.Future Directions:

- Experiment with alternative pretrained architectures (EfficientNet, DenseNet)
- Hyperparameter tuning for batch size, learning rate, and dropout
- Additional augmentation strategies

8.Key Takeaways

- Dataset Size Matters: Smaller datasets overfit easily; increasing training data improves model stability and accuracy.
- CNN from Scratch: Provides decent results but limited by dataset size; overfitting occurs with small datasets.
- Data Augmentation & Dropout: Essential for improving generalization and reducing overfitting.
- Transfer Learning: Pretrained ResNet50 leverages ImageNet features for rapid performance gains, even with limited data.
- Fine-Tuning: Unfreezing top layers allows the model to learn dataset-specific patterns, achieving the highest test accuracy.
- Model Performance Summary:
    - CNN (Scratch, Full dataset): 85.0% test accuracy
    - Pretrained ResNet50 (Fine-tuned): 96.3% test accuracy

9. Conclusion

- CNN from scratch achieved decent accuracy but suffered on small datasets.
- Data augmentation and dropout improved generalization.
- Transfer learning with ResNet50 significantly improved performance.
- Fine-tuning the pretrained model gave the best results.