Università
Ca'Foscari
Venezia

Cats-vs-Dogs Image Classification

Binary image classification with ResNet50.

879779 - Run Jie Xia
27/06/2025

Dataset

Microsoft Cats vs Dogs

Classes: 2 (Cat, Dog)

Size:
25000 images
12500 cats + 12500 dogs

Train-Validation-Test split: 80/10/10
Train set: 10000/class
Validation set: 1250/class
Test set: 1250/class

Dataset Source: [Kaggle]

Resources

- Jupyter notebook
- Key python packages: PyTorch, TorchVision, NumPy, Matplotlib, Seaborn
- Hardware: RTX 4070 SUPER, Ryzen 9 7950X, 64GB RAM

Transfer Learning

A machine learning technique where a pre-trained model, initially designed for one task, is adapted for a related but different task by reusing its learned features.

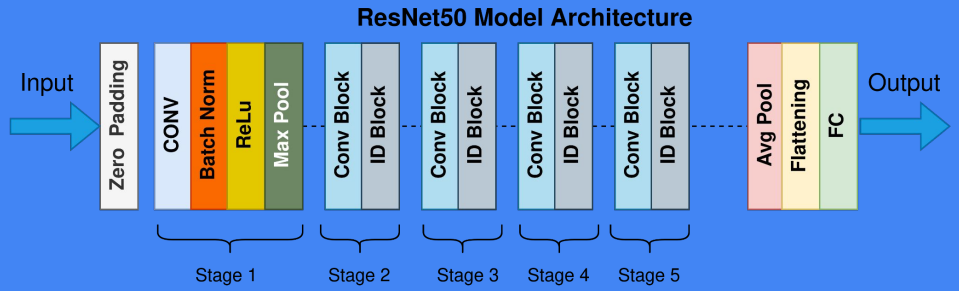
The pre-trained model has already learned to detect basic features (e.g., edges, textures) from a large and diverse dataset. We modify only the last layers of the pre-trained model to specialize it for our specific task.

- Faster training than from scratch
- Requires less data
- Leverages learned feature detectors

Our approach:

- Pre-trained ResNet50 model trained on ImageNet.
- Replaced its original output layer with a custom classification head for binary classification.
- Fine-tuned the last layers of the model to adapt to our dataset.

ResNet50

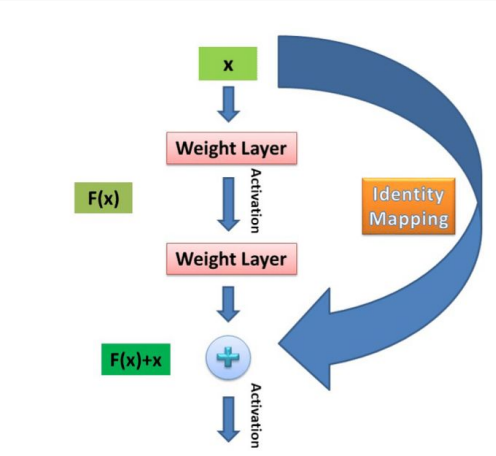


A convolutional neural network (CNN) with 50 layers.
Solves the vanishing gradients problem in very deep networks using residual connections.

Residual Blocks:
Learn residual functions instead of directly mapping inputs to outputs, allowing for deeper networks.

50 Layers:
Composed of convolutional, batch normalization, activation (ReLU), and fully connected layers.

Trained on the ImageNet dataset (1000 classes and millions of images).
Offers learned feature maps that are transferable to new tasks.



ResNet-50 training set up

- Batch size: 32
- Training throughput: 625 batches/epoch
- Loss function: BCE (Binary Cross Entropy)
- Optimizer: Adam (Adaptive Moment Estimation)
- Learning rate: 1e-5 (0.00001)
- Epochs: 30

Freezing Strategy: All ResNet layers frozen
Replacing ResNet50's final layer with a new classification head

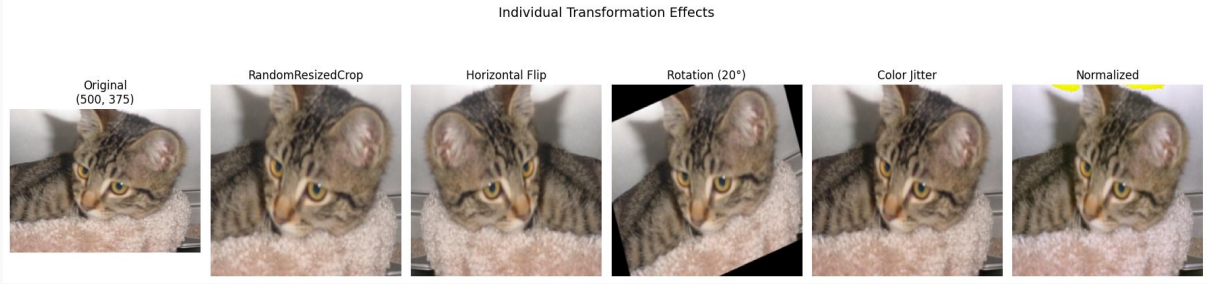
Data Augmentation

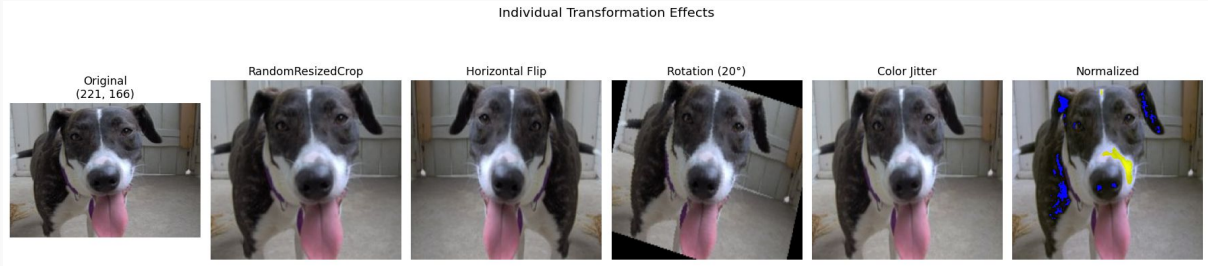
Training Transforms:

- RandomResizedCrop (0.8, 1.0)
- Horizontal Flip
- Rotation (±20°)
- Color Jitter

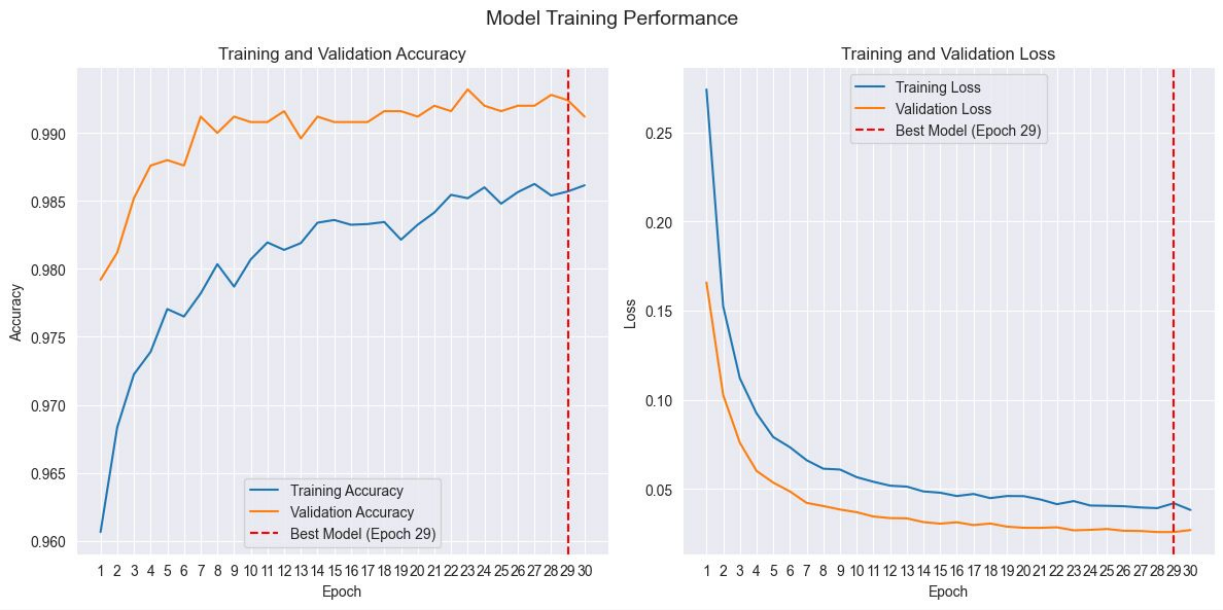
Val/Test Transforms:

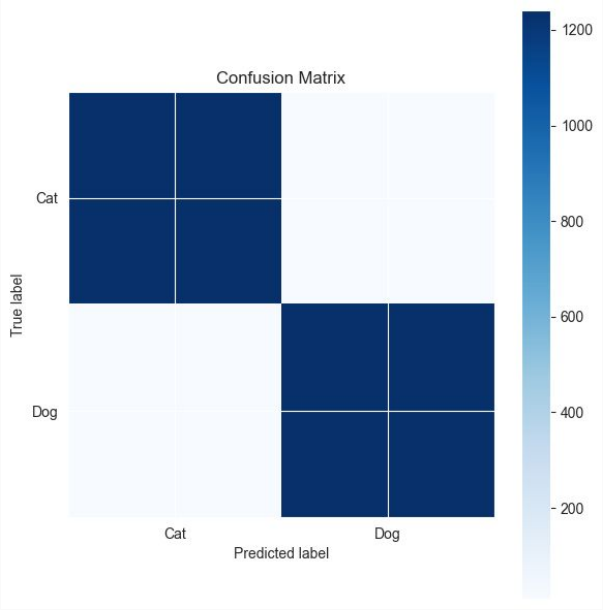
- Resize
- No augmentation to preserve data integrity





Results





Best model:
Epoch 29/30
Train Loss: 0.0420 Acc: 0.9857
Val Loss: 0.0259 Acc: 0.9924

Test Accuracy: 99.04%

	precision	recall	f1-score
Cat	0.99	0.99	0.99
Dog	0.99	0.99	0.99

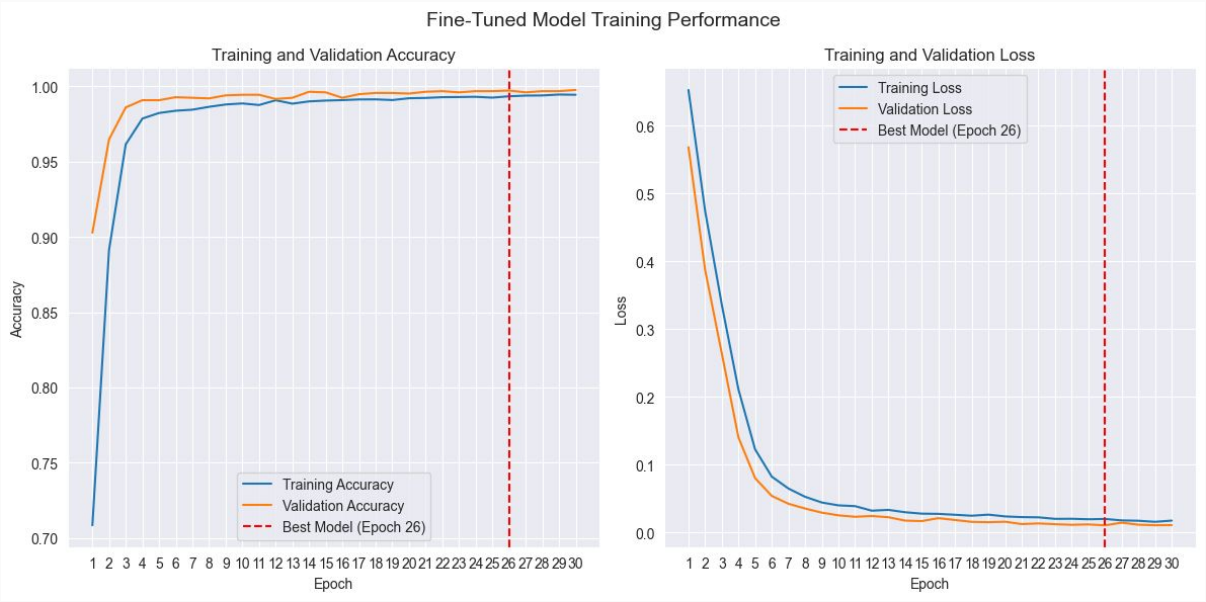
1237	14
10	1241

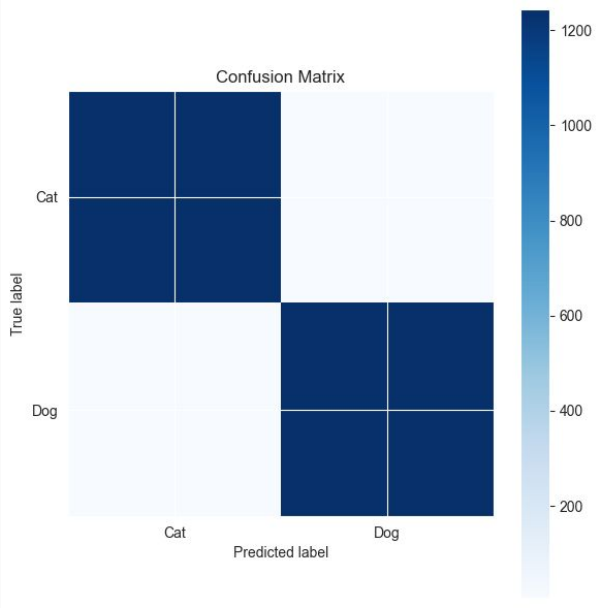
ResNet-50 Fine Tuned training set up

- Batch size: 32
- Training throughput: 625 batches/epoch
- Loss function: BCE (Binary Cross Entropy)
- Optimizer: Adam (Adaptive Moment Estimation)
- Learning rate: 1e-6 (0.000001)
- Epochs: 30

Freezing Strategy: Last 10 ResNet layers unfrozen
Replacing ResNet50's final layer with a new classification head

Results (Fine tuned)





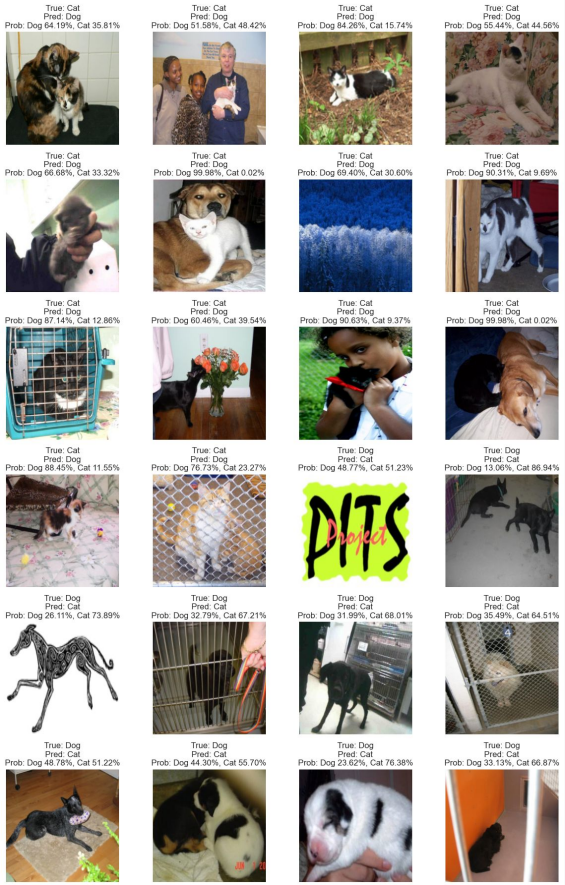
Best model:
Epoch 26/30
Train Loss: 0.0201 Acc: 0.9934
Val Loss: 0.0107 Acc: 0.9972

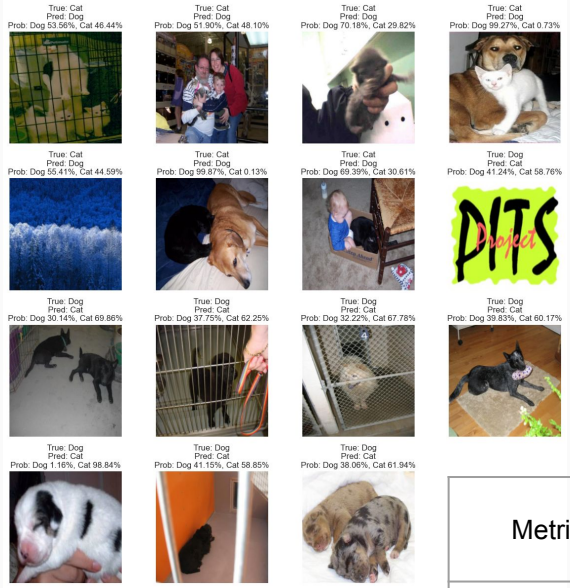
Test Accuracy: 99.40%

	precision	recall	f1-score
Cat	0.99	0.99	0.99
Dog	0.99	0.99	0.99

1244	7
8	1243

Comparison





Metric	Initial	Fine-tuned	Change
Test Accuracy	99.04%	99.40%	+0.36%
Cat Accuracy	98.88%	99.44%	+0.56%
Dog Accuracy	99.20%	99.36%	+0.16%

Conclusions

- Transfer learning is highly effective on this image classification task.
- Fine-tuning improved performance but still has limitations.
- Some misclassifications are due to dataset issues.

Future Work:

- Clean ambiguous/noisy images
- Experiment with unfreezing additional layers or adjusting the dropout rate.

Thank you for your attention