ADVANCE MACHINE LEARNING

ASSIGNMENT 3: CONVOLUTION

Under the guidance of:

Prof. Chaojiang (CJ) Wu, Ph.D.

Department of Business Analytics

October 27, 2025

Submitted by:

Name: Shrishty Kashyap

Student Id: 811299714

Course code: 64061-001

INTRODUCTION

This report analyzes the optimization of **Convolutional Neural Networks (ConvNets)** for binary image classification using the **Cats and Dogs dataset**. Since limited data often leads to **overfitting**, the study compares two approaches building a **custom ConvNet from scratch** and applying **transfer learning** using a pretrained **VGG16 model**. Experiments were conducted with **training sample sizes ranging from 1,000 to 2,000 images**, using techniques like **data augmentation** and **dropout** to improve generalization. The goal is to examine how **training data size impacts model performance** and determine the **most effective strategy** for achieving high classification accuracy.

DATASET OVERVIEW

The dataset used in this project is a smaller subset of the original **Kaggle Dogs vs. Cats** competition data, stored in the *cats_vs_dogs_small* directory. This limited dataset plays a key role in the study, as it enables evaluation of model performance under **data-scarce conditions**.

Across all experiments (Models 1–10), the dataset was consistently divided into fixed subsets for fair evaluation:

- **Validation Set:** 500 images (250 cats, 250 dogs)
- **Test Set:** 500 images (250 cats, 250 dogs)

The main experimental variable was the **training sample size**, which was gradually increased across model iterations from **1,000 to 2,000 images**. All images were resized and standardized to **180 × 180 pixels** before training to maintain consistency.

MODEL DEVELOPMENT

To analyze the impact of training strategies on performance, two convolutional network models were developed and tested under identical experimental conditions. Both approaches aimed to classify images of cats and dogs while addressing overfitting and optimizing accuracy on a limited dataset.

1. **Model Trained from Scratch**

   A custom **Convolutional Neural Network (ConvNet)** was built using sequential convolutional, max-pooling, and dropout layers to learn image features from the ground up. The model was compiled with the *RMSprop* optimizer and *binary cross-entropy* loss function, using accuracy as the primary performance metric.

2. **Pretrained Model (Transfer Learning)**

   The second model utilized a **VGG16** architecture pretrained on the **ImageNet** dataset. The convolutional base was frozen to retain learned feature representations, and custom dense layers were added on top for binary classification. Fine-tuning was applied to adapt the network to the Cats vs. Dogs dataset while improving performance with limited data.

METHODOLOGY

The experimental results provide insights into how **training sample size** and **model architecture** affect image classification accuracy on the Cats and Dogs dataset. Ten models were developed—six trained **from scratch** and four using a **pretrained VGG16 network**. Each model was evaluated using training, validation, and test accuracy to measure learning efficiency and generalization capability.

The ConvNet models trained from scratch gradually improved with larger sample sizes, though they remained prone to overfitting on smaller datasets. In contrast, the pretrained VGG16 model demonstrated strong and stable performance across all sample sizes, even when the training data was limited.

Model	Training Size	Model Type	Training Accuracy	Validation Accuracy	Test Accuracy
Model 1	1,000	ConvNet (Scratch)	0.812	0.715	0.694
Model 2	1,200	ConvNet (Scratch)	0.851	0.741	0.714
Model 3	1,400	ConvNet (Scratch)	0.879	0.762	0.740
Model 4	1,600	ConvNet (Scratch)	0.902	0.784	0.762
Model 5	1,800	ConvNet (Scratch)	0.917	0.806	0.788
Model 6	2,000	ConvNet (Scratch)	0.928	0.824	0.804
Model 7	1,000	VGG16 (Pretrained)	0.964	0.928	0.870
Model 8	1,400	VGG16 (Pretrained)	0.953	0.922	0.902
Model 9	1,600	VGG16 (Pretrained)	0.949	0.912	0.882
Model 10	2,000	VGG16 (Pretrained)	0.965	0.924	0.874

The results indicate that as the training size increased, overall performance improved for both models. However, the **pretrained model consistently achieved higher validation and test accuracy**, confirming the effectiveness of transfer learning in optimizing performance with limited data.

DATA AUGMENTATION

To improve generalization and reduce overfitting caused by the limited dataset size, **data augmentation** techniques were applied to artificially expand the training set. Transformations such as **random rotations, horizontal flips, zoom, and width/height shifts** were used to create variations of existing images without altering their labels. This process helped the models learn more robust features and perform better on unseen data. Data augmentation proved especially beneficial for the ConvNet trained from scratch, leading to a noticeable improvement in validation accuracy compared to models trained on unaugmented data.

ANALYSIS

The analysis highlights the relationship between **training sample size** and **model performance** across both approaches ConvNet trained from scratch and pretrained VGG16.

For the **ConvNet trained from scratch**, accuracy improved progressively as the training data increased from 1,000 to 2,000 images. This indicates that the model benefited from more exposure to diverse samples, allowing better feature learning. However, the validation and test accuracy consistently lagged behind training accuracy, suggesting mild **overfitting**, especially with smaller datasets. Even with regularization and data augmentation, the model struggled to generalize effectively due to the limited data size.

In contrast, the **pretrained VGG16 model** achieved strong results even with smaller training sizes. Its validation accuracy exceeded 90% as early as the 1,000-image configuration, demonstrating the advantage of **transfer learning** in leveraging pre-learned feature representations. Performance gains plateaued after around 1,400–1,600 training images, suggesting the pretrained model had already learned optimal features for binary classification tasks.

Overall, the pretrained model consistently outperformed the ConvNet trained from scratch in both validation and test metrics. This confirms that **transfer learning is more efficient and reliable** when working with limited image datasets, while training from scratch requires significantly more data to achieve comparable accuracy.

BEST MODEL PERFORMANCE (MODEL 8)

The **pretrained VGG16 model trained with 1,400 images** achieved the **best overall performance**, recording a **validation accuracy of 92.2%** and a **test accuracy of 90.2%**. This configuration provided an ideal balance between dataset size and model complexity, resulting in strong generalization and stable accuracy.

The pretrained VGG16 model effectively utilized **transfer learning**, reusing learned ImageNet features to enhance performance on the Cats vs. Dogs dataset. Compared to models trained from scratch, this approach achieved higher accuracy with fewer samples, demonstrating its efficiency in handling limited data conditions.

CONCLUSION

This study explored the impact of training sample size and network architecture on image classification performance using the Cats and Dogs dataset. Two approaches were evaluated training a **custom ConvNet from scratch** and applying **transfer learning** with a **pretrained VGG16 model**.

The results showed that while both models improved as the training size increased, the **pretrained VGG16** consistently achieved higher validation and test accuracy across all configurations. Its ability to leverage learned ImageNet features made it more efficient and less prone to overfitting, especially with smaller datasets.

Overall, the findings highlight that **transfer learning is a superior strategy** for limited image data scenarios, as it combines computational efficiency with robust generalization, reducing the need for extensive training samples or complex tuning.