Summary Report

Assignment 3

Advanced Machine Learning_64061

MSBA

Submitted by: Dileep Kumar Pasala

❖ Introduction:

Computer vision has been transformed by deep learning, which makes it possible for machines to recognize and classify images with amazing accuracy. The Convolutional Neural Network (CNN), a potent design that automatically learns to extract significant spatial information from visual data using stacked convolution and pooling layers, is at the core of this metamorphosis.

This study employs two different learning algorithms to classify photos of dogs and cats. CNNs are trained from scratch, and learning behavior is observed using different dataset sizes. Utilizing pretrained ImageNet weights, transfer learning with the VGG16 architecture improves accuracy and generalization

The objective was to ascertain whether pretrained architecture offers quantifiable benefits over scratch-built models and to examine how model complexity and dataset size affect performance.

❖ Problem statement:

The goal is binary classification, which involves correctly classifying pictures as either "dog" or "cat." In order to enhance classification performance and reduce overfitting, CNN architecture must be implemented and optimized. The study attempts to identify the best strategy for attaining high accuracy by experimenting with various training sample sizes and using both training from scratch and pre-trained models. To improve the generalization and robustness of the model, methods like data augmentation and regularization will be used.

❖ Primary Objective:

1. Examine how the size of the training sample affects CNN's ability to classify images.
2. Examine the differences between utilizing a pre-trained network and building a CNN from scratch for the purpose of identifying photos of dogs and cats.
3. To lessen overfitting and enhance model generalization, investigate optimization strategies including data augmentation and regularization.
4. Determine the ideal training sample size that produces the most accurate predictions.
5. Examine the trade-offs in deep learning-based image categorization between model performance, data availability, and computational cost.

❖ Overview of the Dataset:

- Kaggle Cats vs. Dogs dataset
- About 12,475 cats and 12,480 dogs were included in the total cleaned images.
- Picture Dimensions: 180 x 180 pixels
- 500–3,500 photos every class for training.
- Validation: 500 pictures
- Test: 500 pictures

Data cleaning included assuring balanced class representation, eliminating corrupt data, and normalizing channels to RGB.

❖ Design of Models:

✓ ***CNN Models (1–10):***

- trained on various dataset sizes from the ground up.
- Architecture: dropout → sigmoid → dense → pooling → sequential convolution.
- Regularization: Data augmentation (flip, rotation, zoom) and dropout (0.5).

✓ ***VGG Transfer Learning Models (1–3):***

- used a VGG16 convolutional foundation that has been pretrained.
- Head of custom classification (Dense + Dropout).
- adjusted the number of top layers to suit different dataset sizes.

❖ Techniques for Augmenting Data:

Type	**Transformation**
Random Flip	Horizontal
Random Rotation	±10° – 15°
Random Zoom	20–25%

These augmentations prevented overfitting and improved generalization across smaller datasets.

- **Overall Model Performance Summary:**

Model	*Training Images per Class*	*Test Accuracy*	*Test Loss*	*Rank (by Accuracy)*
VGG Model 3	**500**	**0.966**	**0.4927**	**1**
VGG Model 2	3500	0.908	0.2149	2
Model 8	3000	0.878	0.295	3
Model 9	3000	0.87	0.2909	4
Model 7	3000	0.838	0.4262	5
Model 10	3500	0.828	0.3491	6
Model 5	2500	0.814	0.3967	7
Model 6	2500	0.742	0.4982	8
Model 4	500	0.714	0.5509	9
Model 2	500	0.688	0.5758	10
Model 3	500	0.652	0.6253	11
Model 1	500	0.65	0.6446	12
VGG Model 1	3000	0.518	0.7124	13

- **Performance Visualization:**

  a. Deeper CNNs and pretrained models show reduced loss and higher accuracy.

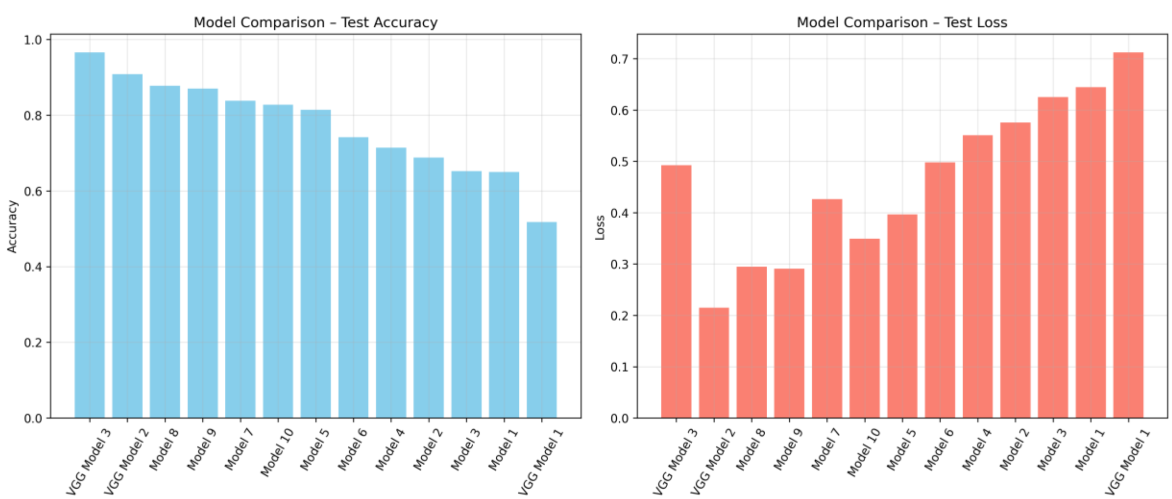


Figure 1. Comparison of test accuracy and losses for all models.

b. The scatter plot confirms the value of transfer learning by showing that VGG-based models cluster in the high-accuracy, low-loss region.

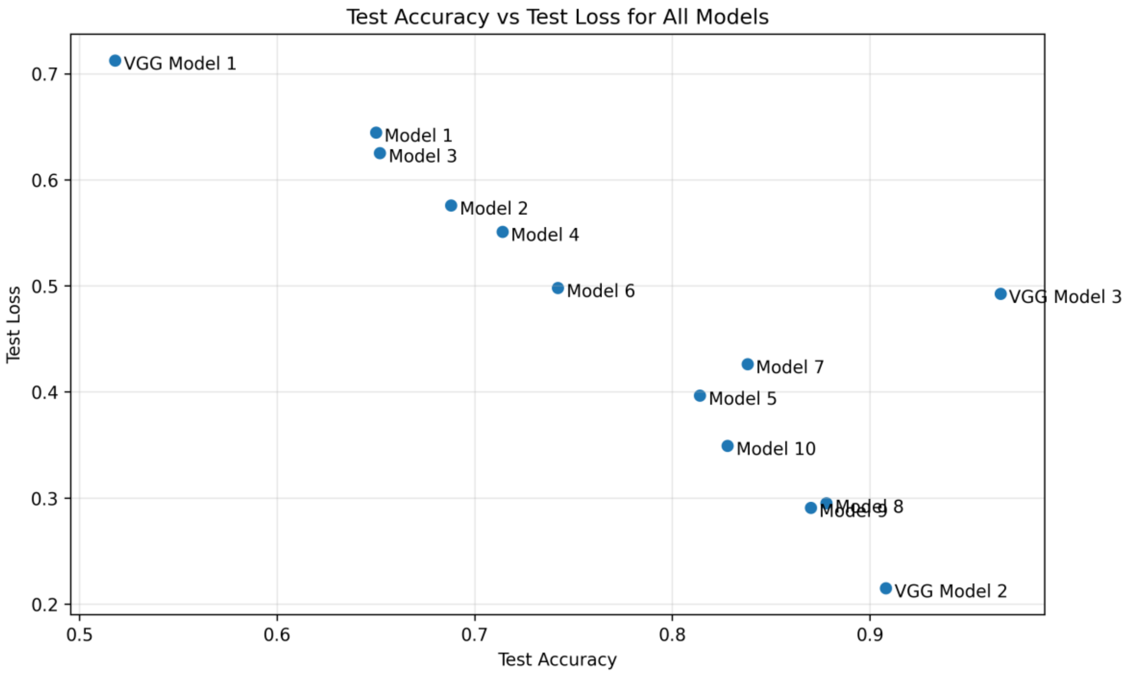


Figure 2. Scatter plot of test accuracy vs test loss for all models.

c. Model accuracy grows with dataset growth until it reaches a plateau at about 3,000 photos per class. Regardless of the amount of the dataset, VGG models are consistently better.

Figure 3. Impact of training images per class on test accuracy for CNN and VGG models.

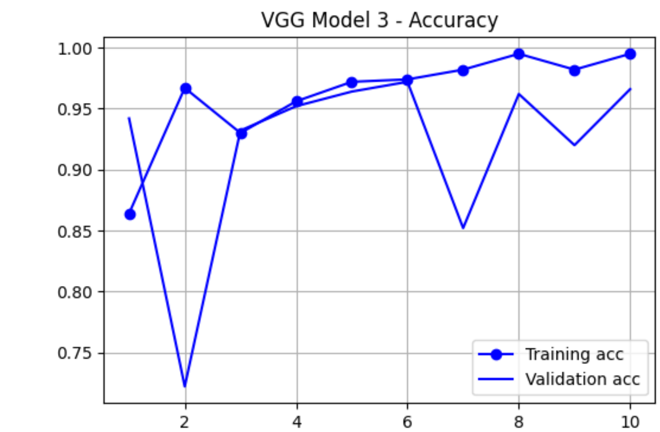
❖ Detailed Model Comparison:

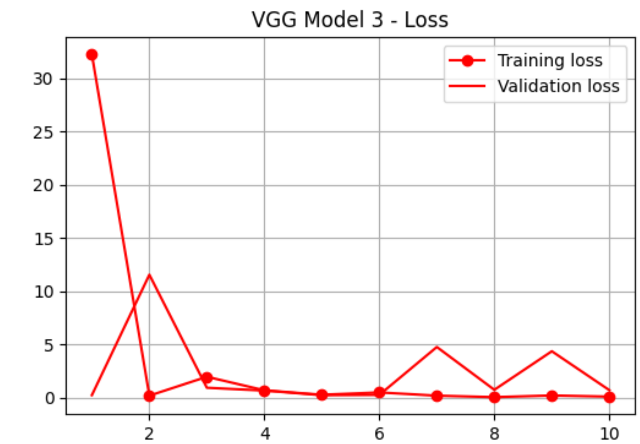
CNN Models

Model	Training Images per Class	Test Accuracy	Test Loss	Rank (by Accuracy)
Model 8	3000	0.878	0.295	3
Model 9	3000	0.87	0.2909	4
Model 7	3000	0.838	0.4262	5
Model 10	3500	0.828	0.3491	6
Model 5	2500	0.814	0.3967	7
Model 6	2500	0.742	0.4982	8
Model 4	500	0.714	0.5509	9
Model 2	500	0.688	0.5758	10
Model 3	500	0.652	0.6253	11
Model 1	500	0.65	0.6446	12

VGG Transfer Learning Models

Model	Training Images per Class	Test Accuracy	Test Loss	Rank (by Accuracy)
VGG Model 3	500	0.966	0.4927	1
VGG Model 2	3500	0.908	0.2149	2
VGG Model 1	3000	0.518	0.7124	13



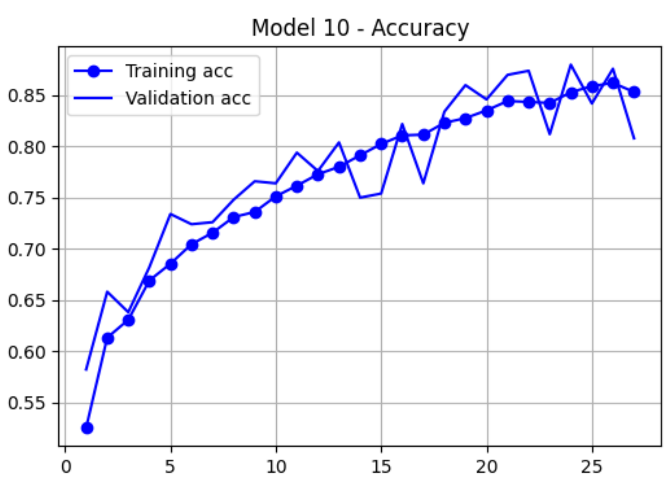


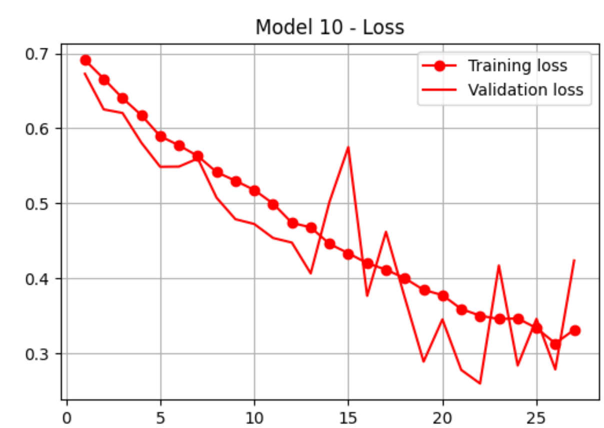

```
16/16 ━━━━━━━━━━━━━━━━━━━━ 2s 52ms/step - accuracy: 0.9697 - loss: 0.738
VGG Model 3 → Test Accuracy: 0.966 | Test Loss: 0.493

✅ VGG Model 3 Final Results:
Test Loss: 0.493
Test Accuracy: 0.966
```

VGG16 Model3:

Test Loss: 0.493

Test Accuracy: 0.966






```
16/16 ━━━━━━━━━━━━━━━━━━━━ 0s 7ms/step - accuracy: 0.8433 - loss: 0.2822
Model 10 → Test Accuracy: 0.828 | Test Loss: 0.349

✅ Model 10 Final Results:
Test Loss: 0.349
Test Accuracy: 0.828
```

CNN Model 10:

Test Loss: 0.349

Test Accuracy: 0.828

Observations:

- CNNs from scratch showed steady accuracy growth as the dataset increased, emphasizing the importance of sufficient data.
- Data augmentation successfully reduced overfitting in smaller datasets, improving validation stability.
- Transfer learning using VGG16 achieved up to 96.6% accuracy, even with smaller training sets, confirming its strong feature extraction capability.
- The best performing model overall is VGG Model 3, achieving a test accuracy of **0.966** and test loss of **0.493**.
- Transfer-learning (VGG-based) models consistently outperform all models trained from scratch, demonstrating the significant advantage of leveraging pre-trained ImageNet features even with limited data.
- Accuracy generally increases with training sample size, with performance gains becoming less pronounced beyond approximately 3 000 images per class.

Conclusion

This work highlights the effectiveness of transfer learning for image classification tasks, especially when the available dataset is relatively small. Using a pretrained convolutional backbone such as VGG16 yields significantly better generalization and training stability compared to CNNs trained entirely from scratch. Further improvements could potentially be explored by fine-tuning more layers of the VGG base or experimenting with other pre-trained architecture.

Important lessons learned:

- Up to about 3,000 photos per class, accuracy increases with dataset size.
- Small-data learning is greatly enhanced by data supplementation.
- A significant increase in accuracy (≈+10–15%) is provided via transfer learning.
- The best generalization and accuracy (96.6%) were attained by VGG Model 3.
- Future improvements might involve experimenting with ResNet50, InceptionV3, or optimizing VGG16's deeper layers for additional accuracy gains.