Assignment 3: Convolution

Nishanth Gurujala

INTRODUCTION

This report discusses how Convolutional Neural Networks (CNNs) can be utilized in classifying binary images, that is, classifying cat and dog images. The primary aim was to evaluate the performance of models that are trained directly against models that are based on pretrained archives and to change the size of the training set. We started with a small training sample, which was gradually enlarged to test the effect that this variation had on model accuracy, generalization, and robustness. Finally, it was aimed at identifying the best training strategy, be it scratch trained or pretrained, in order to reach high performance levels when training on various data scales.

DATA OVERVIEW

The information that will be utilized in the study is the famous Cats and Dogs dataset which includes the information of the thousands of images of two classes: Cats and Dogs. A sub-set of this data was utilized as the experimental consistency. All the images were preprocessed to a standardized format as required by CNN inputs. The sample size of training was different in experiments, and the validation and test sets were constant to make the performance assessment fair and similar.

DATA PREPROCESSING

- Accessing the image files.
- Image decoder: A decoder that transforms the JPEG files into the pixel grids of RGB format.
- Normalization: Transforming pixel values of the original range of [0, 255] to [0, 1], to enhance numerical stability and training performance.

MODEL TRAINING

- CNN trained: A custom CNN trained on subsets of the data of varying sample sizes.
- Pretrained CNN: This is a transfer learning model that uses pretrained weights to improve feature extraction and minimize computation cost.
- Both models were streamlined with different sample size so as to test the effects of quantity of data on the performance.

PERFORMANCE EVALUATION

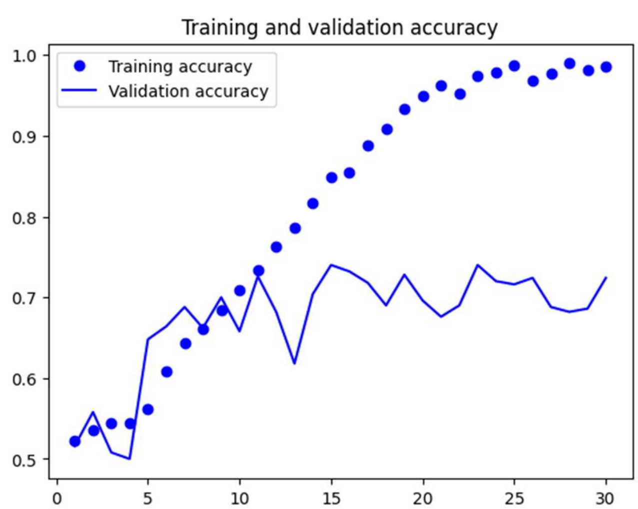
The accuracy and loss measures were used to evaluate the performance of the model on training and validation sets. The accuracy of validation was monitored closely to identify possible overfitting and also assess the capability of the model to work with unseen data.

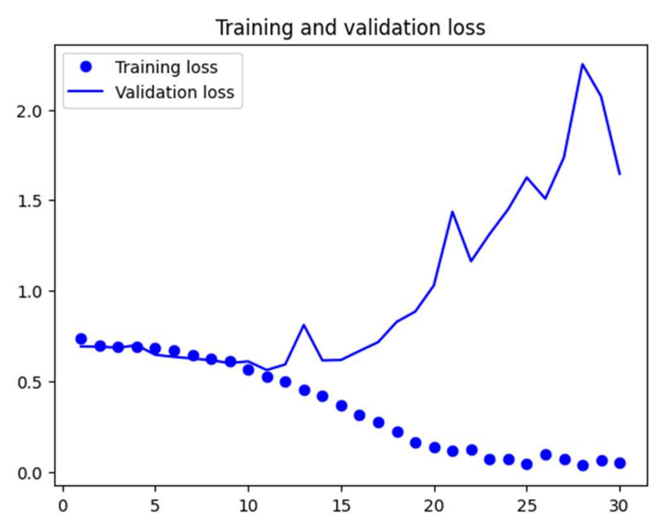
PERFORMANCE ANALYSIS AND METHODOLOGY.

Consistent CNN architecture has been adopted to evaluate the effect of training sample size on the model performance with experiment setups (n=5) varying in the size of the training data (1,000, 1,400, 1,600, 1,800, and 2,000 images). The test and validation set was shared across all the models.

Both CNNs had five convolutional layers with ReLU activation functions, and a dense final layer with sigmoid activation to classify two categories. The models were summarized using the binary cross-entropy, RMSprop optimizer, and accuracy as the evaluation of accuracy. The consistency allowed it to guarantee that the differences in the observed performances could be attributed to the size of the training data and not the architectural or optimization alterations.

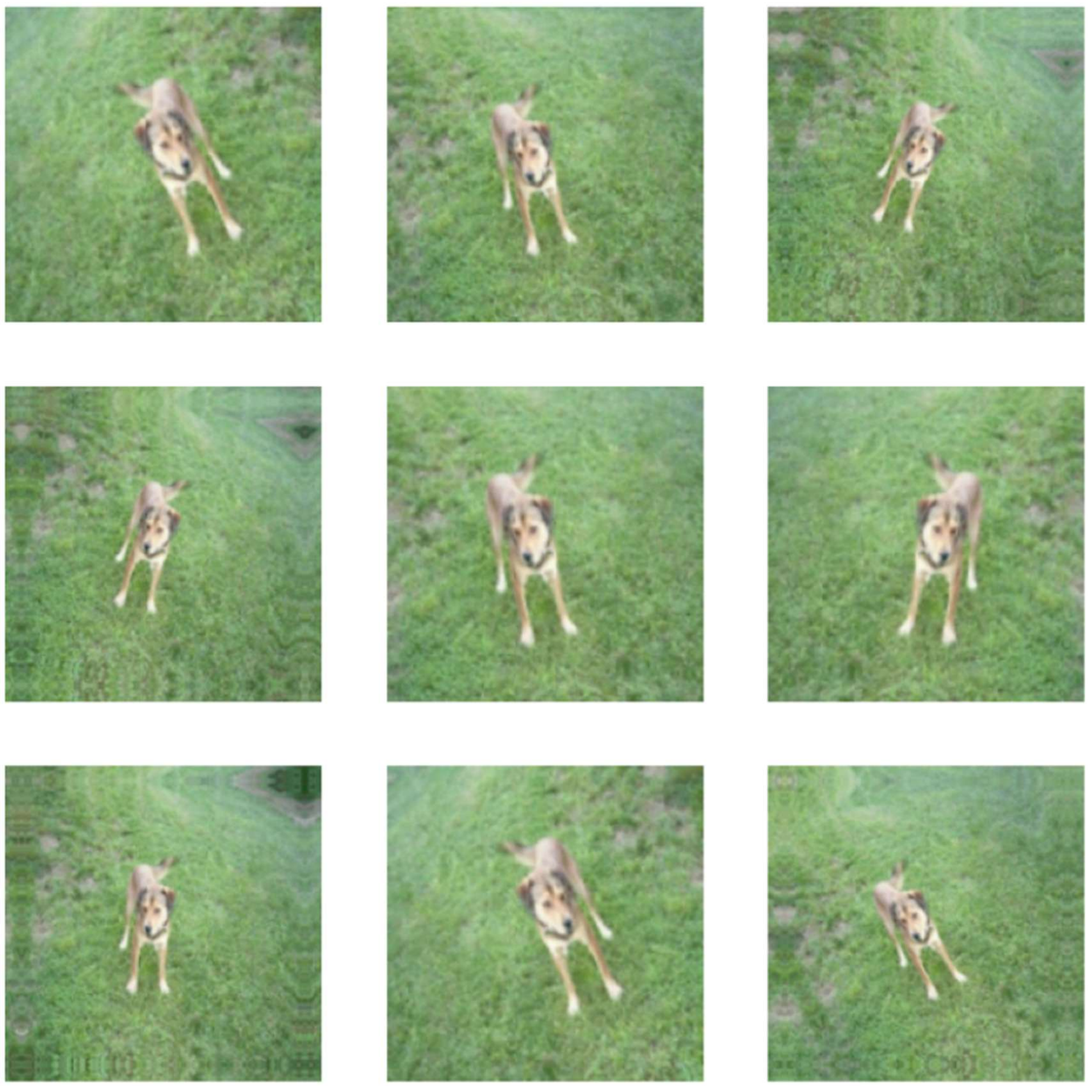
In order to promote generalization, random flipping, rotation, and zooming data augmentation methods were used to provide variability in the training data. Also, the transfer learning using the pretrained CNN was applied to use the learned features and enhance the accuracy and training efficiency, especially with smaller datasets.Each of the models was trained with the depth of 30 epochs, which was a compromise between the learning and computational ability. This approach allowed us to comparatively investigate the effects of the size of data, data augmentation, and transfer learning on the accuracy, loss, and the overall robustness of models.





DATA AUGMENTATION

Data augmentation was also used to enhance the performance of the models particularly where the training data was small. The new image variations were generated by the application of random transformation- horizontal flips, rotations, and zooms on the existing data. This procedure exposed the model to various speech and views and it became more generalized to see unseen pictures. Consequently, data augmentation was influential in the enhancement of dataset diversity, reduction of overfitting and enhance the overall classification accuracy.



Analysis

Model and Sample size	Method	Validation Accuracy	Test Accuracy	Validation Loss	Test Loss
Model 1 Training – 1000 Validation – 500 Test- 500	Without Augmentation	67.60%	64.80%	0.588	0.625
	With Augmentation	77.60%	70.40%	0.471	0.55
Model 2 Training – 1400 Validation – 500 Test- 500	Without Augmentation	71.00%	70.60%	0.561	0.586
	With Augmentation	80.40%	77.00%	0.426	0.545
Model 3 Training – 1600 Validation – 500 Test- 500	Without Augmentation	72.60%	69.40%	0.562	0.577
	With Augmentation	82.00%	82.20%	0.415	0.453
Model 4 Training – 1800 Validation – 500 Test- 500	Without Augmentation	70.00%	68.40%	0.572	0.629
	With Augmentation	75.80%	75.00%	0.485	0.546
Model 5 Training – 2000 Validation – 500 Test- 500	Without Augmentation	72.00%	71.60%	0.547	0.579
	With Augmentation	84.60%	79.80%	0.361	0.476

PRETRAINED MODELS TABLE

Model traning sample size	Method	Validation Accuracy	Test Accuracy	Validation Loss	Test Loss
Model 6- 1000	With Augmentation	97.80%	97.00%	2.456	5.307
Model 7 - 1400	With Augmentation	98.60%	96.00%	1.339	6.235
Model 8 – 1600	With Augmentation	98.20%	98.00%	2.713	2.201
Model 9 - 1800	With Augmentation	98.80%	97.80%	0.84	3.394
Model 9 - 1800	With Augmentation	98.60%	97.80%	1.32	3.427

TRAINING FROM SCRATCH & PRETRAINED MODEL

Model	Method	Validation Accuracy	Test Accuracy	Validation Loss	Test Loss
Model 3 - 1600	With Augmentation	82.00%	82.20%	0.415	0.453
Model 9 - 1800	With Augmentation	98.80%	97.80%	0.84	3.394

CONCLUSION

Among the models trained from scratch, **Model 3 (trained with 1,600 samples and data augmentation)** demonstrated the **best overall performance**, achieving a **test accuracy of 82.2%** and a **low-test loss of 0.453**, indicating strong generalization to unseen data. Data augmentation played a critical role in improving model robustness by introducing greater image diversity and preventing overfitting. Model 3 (trained from scratch with 1,600 samples and augmentation) achieved the best balance of accuracy and loss, showing strong generalization. Among pretrained models, Model 8 (1,600 samples with augmentation) performed best, reaching 98% test accuracy with low loss. Model 5 (2,000 samples with augmentation) provided reliable generalization, while Model 9 (1,800 samples) achieved higher accuracy but showed signs of overfitting. Overall, convolutional neural networks (ConvNets) proved effective even on small datasets, but overfitting remains a challenge. Data augmentation and pretrained models help address this by improving robustness and efficiency, highlighting the trade-off between training from scratch and transfer learning depending on dataset size and computational resources.