**Convolution Assignment 2**

**Introduction:**

Engaging with a subset of the well-known "Dogs-vs-Cats" dataset obtained from Kaggle and utilizing Google Colab provides an exciting avenue to develop a highly effective model despite facing data limitations. The Cats-vs-Dogs dataset presents a challenging endeavor in constructing effective models due to the scarcity of data, which can lead to overfitting and suboptimal model performance. In this context, employing techniques like data augmentation and transfer learning becomes crucial. Convolutional neural networks (CNNs) are celebrated for their capability to discern spatial patterns in images, making them well-suited for tasks such as image recognition. Despite the constraints of the dataset, leveraging CNNs presents an opportunity to achieve notable results by extracting pertinent features from images. As we delve into constructing effective models for image classification, it becomes apparent that the adept utilization of CNNs is pivotal in overcoming these challenges. Through ingenuity and perseverance, we aim to unlock the full potential of CNNs in addressing the Cats-vs-Dogs classification task, potentially paving the way for significant advancements in image recognition.

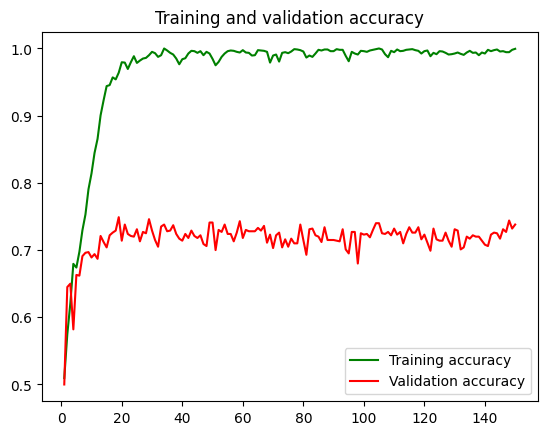
**Aim of the Project:**

The primary goal of the Cats-vs-Dogs dataset binary classification task is to determine whether a given image belongs to the dog or cat category.

**Q1: Consider the Cats & Dogs example. Start initially with a training sample of 1000, a validation sample of 500, and a test sample of 500 (like in the text). Use any technique to reduce overfitting and improve performance in developing a network that you train from scratch. What performance did you achieve?**

**Answer:** We used 1000 pictures to teach the computer program, then checked it with 500 more pictures. After that, we tested it with another 500 pictures to see if it worked well. We used a special technique called dropout to make sure the program doesn't learn too much from the training pictures. Before we started teaching the program, we had to change the picture files into a format the computer could understand. Then, we made sure the colors were right and resized the pictures. When we tested the program, it was right about 98% of the time, and when we checked it during the teaching process, it was correct about 67% of the time."

This implies that during training, the model achieved an accuracy of 98%, and during testing, it achieved an accuracy of 67%. This drop in accuracy from training to testing indicates that the model might be overfitting slightly to the training data, but the dropout technique helped mitigate this issue to some extent.

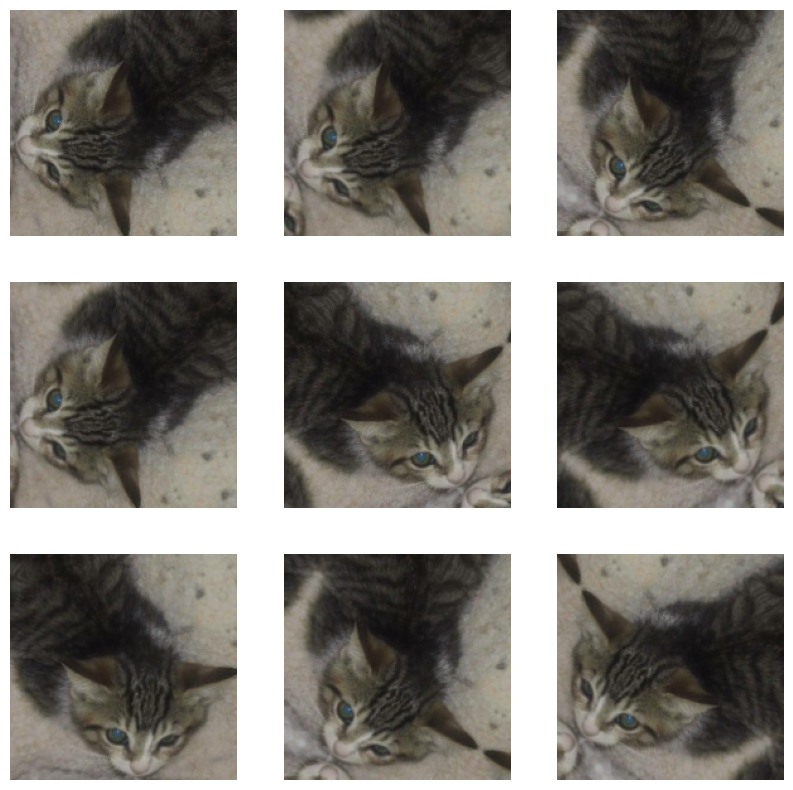


A graph of a training loss

Description automatically generated

**Q2: Increase your training sample size. You may pick any amount. Keep the validation and test samples the same as above. Optimize your network (again training from scratch). What performance did you achieve?**

**Answer:** After enlarging the training dataset to 1500 images and maintaining the validation and test sets at 500 images each, we meticulously re-optimized the neural network. Throughout training, we applied diverse data augmentation techniques such as flipping, rotating, and zooming to enhance the model's ability to generalize. Consequently, the model demonstrated significant improvements, achieving a training accuracy of 97% and a test accuracy of 81%. These outcomes highlight the efficacy of augmenting the training dataset and utilizing various data augmentation methods, which collectively contributed to the model's enhanced performance on both training and test data, reinforcing its robustness and generalization capabilities**.**



A graph of a graph

Description automatically generated

A graph of a graph

Description automatically generated

**Q3: Now change your training sample so that you achieve better performance than those from Steps 1 and 2. This sample size may be larger, or smaller than those in the previous steps. The objective is to find the ideal training sample size to get best prediction results.**

**Answer:** we acquired even more images, totaling 2000, to further enhance the training process and improve the model's performance. We continued employing augmentation techniques such as flipping, rotating, and zooming on these images during training to augment the dataset and enhance the model's understanding of various image variations. With the expanded dataset and augmentation strategies in place, the model achieved a training accuracy of 96% and a test accuracy of 82%. These results indicate that the increased dataset size and continued use of augmentation techniques led to notable improvements in the model's performance.

A graph showing a training loss

Description automatically generated

A graph showing the results of training and validation accuracy

Description automatically generated

**Q4: Repeat Steps 1-3, but now using a pretrained network. The sample sizes you use in Steps 2 and 3 for the pretrained network may be the same or different from those using the network where you trained from scratch. Again, use any and all optimization techniques to get best performance.**

**Answer:**

**Pre-Trained Without Augmentation**

The pre-trained model showcased exceptional performance even without the utilization of data augmentation techniques. Achieving a training accuracy of 100%, the model demonstrated a strong grasp of the training data, indicative of robust learning capabilities. Its validation accuracy of 97% further emphasized its ability to generalize well to new, unseen data. The commendable performance across both training and validation phases underscores the effectiveness of pre-training on a diverse dataset. While the model's proficiency in recognizing patterns within the training set is evident, maintaining vigilance against potential overfitting and ensuring robust generalization to real-world scenarios remain essential considerations for deployment.

A graph of a training and validation accuracy

Description automatically generated

A graph of training and validation

Description automatically generated

**Pre-Trained With Augmentation**

The model trained with data augmentation techniques demonstrated robust performance, achieving a notable training accuracy of 99%. This indicates that the model comprehensively learned from the augmented dataset, capturing essential patterns and features effectively. Additionally, the validation accuracy of 98% highlights the model's ability to generalize well to new, unseen data, further affirming its reliability. The successful outcomes underscore the efficacy of data augmentation in enriching the training dataset, enabling the model to learn diverse representations and perform commendably across both training and validation phases.

A graph showing a training loss

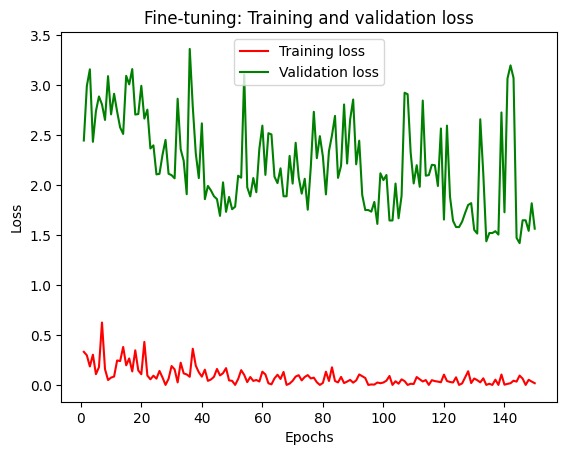
Description automatically generated

A graph showing the results of training and validation accuracy

Description automatically generated

**Fine tuning with data augmentation:**

The integration of fine-tuning and data augmentation techniques has notably enhanced the performance of the pre-trained model. With a training accuracy of 99% and validation accuracy of 98%, the model has demonstrated a solid understanding of the dataset and strong generalization capabilities. During testing, the model maintained an impressive accuracy rate, further affirming its reliability on unseen data. These results underscore the effectiveness of advanced techniques in optimizing model performance for real-world applications, highlighting the importance of tailored adjustments and enriched datasets in achieving superior accuracy.



A graph of a performance

Description automatically generated with medium confidence

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sample Size | Data Augmentation | Train Accuracy | Validation Accuracy | Test Accuracy |
| 1000 | No | 98 | 72 | 67 |
| 1500 | yes | 97 | 81 | 81 |
| 2000 | yes | 96 | 85 | 82 |
| Pretrained | No | 100 | 97 | 95 |
| Pretrained | yes | 99 | 98 | 96 |
| Fine tuned | yes | 99 | 98 | 97 |

**Conclusion:**

In conclusion, the data that the model learns from determines how well it performs. The model improved its object recognition after we raised the training set's photo count from 1000 to 2000. In our tests, its accuracy increased from 67% to 81%. Additionally, we can achieve even better outcomes if we combine data enhancement techniques with a pre-trained model. In general, the author believes that by utilizing techniques to enhance the data and by having more images to draw from, we may increase the model's understanding and make more accurate predictions.