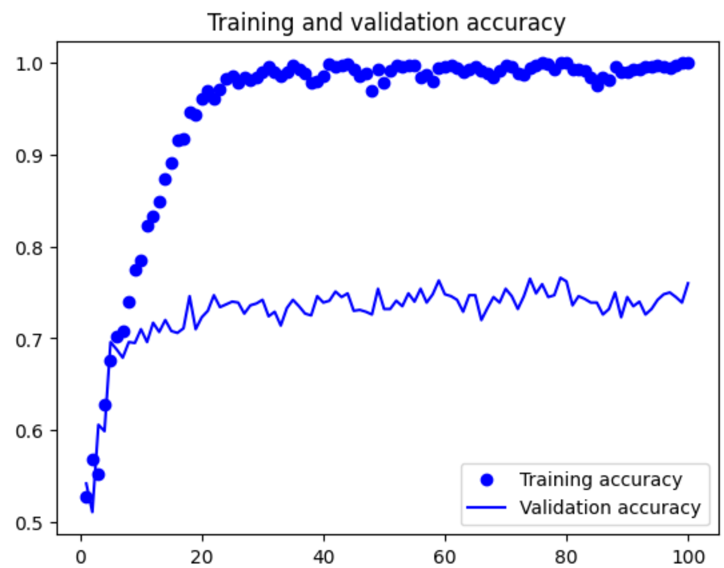
Analysis of the Correlation Between Training Sample Size and Selection of Neural Network for Image Categorization

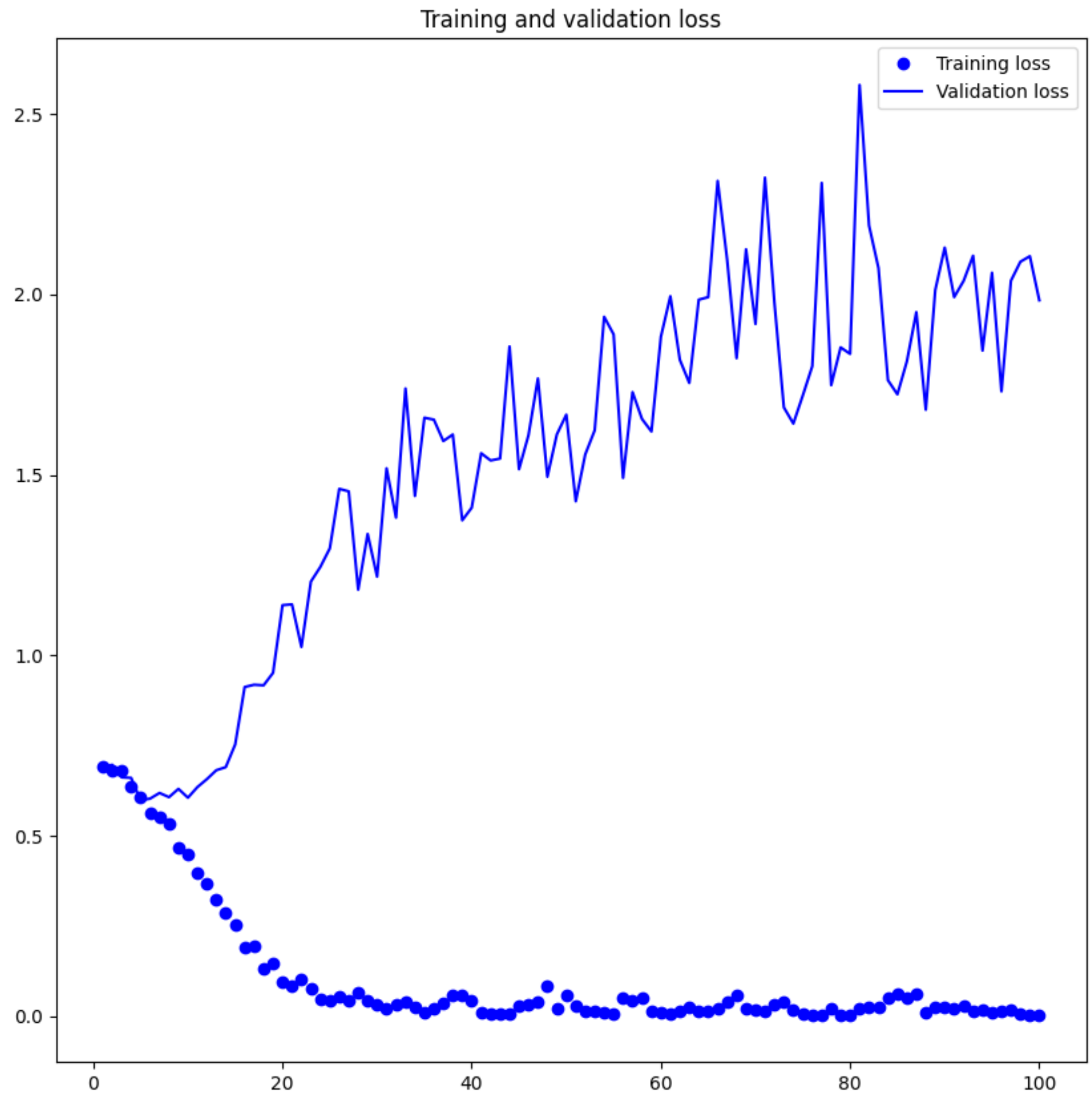
This document discusses the creation of a specialized software known as a convolutional neural network (CNN). This software is designed to identify whether a given image contains a cat or a dog. The images utilized for training the software are sourced from Kaggle. Although numerous images are available, only a subset of 2000 images is employed for training the program.

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 you train from scratch. What performance did you achieve?

Answer:

Initially, a dataset comprising 1000 images was employed for training the computational model, followed by an additional evaluation using 500 supplementary images. Subsequently, another set of 500 images was utilized for comprehensive testing of the model's efficacy. To prevent overfitting, a specialized method known as dropout was implemented to constrain the program's reliance on the training dataset. Preprocessing steps involved converting the image files into a machine-readable format and ensuring color accuracy and image resizing. During testing, the program demonstrated an accuracy rate of approximately 99.95%, whereas its performance during the training phase indicated an accuracy of around 76%.



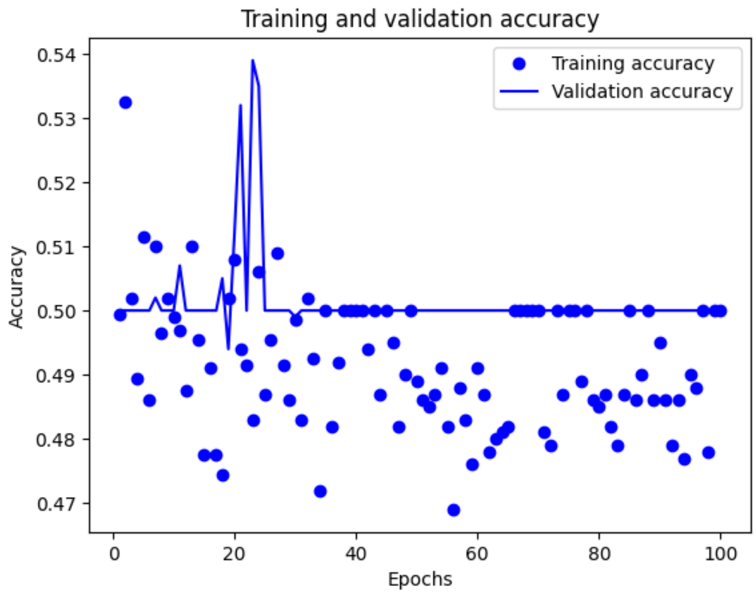


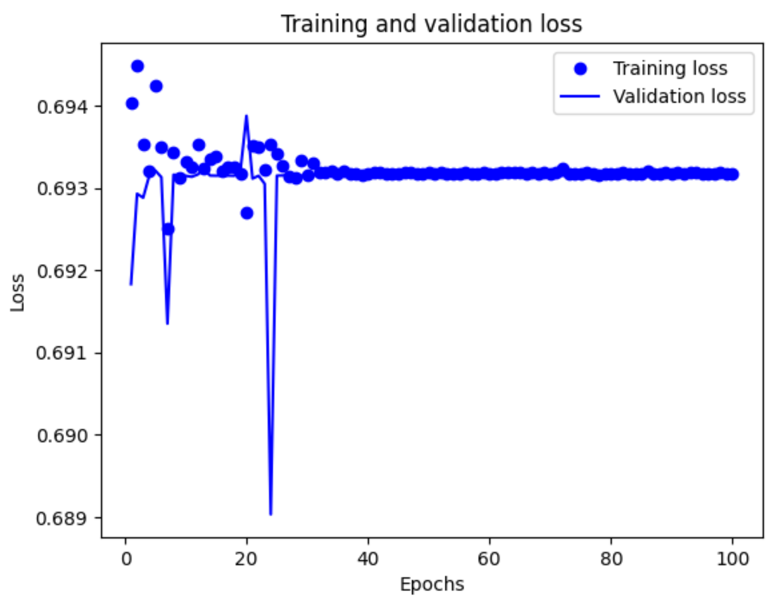
Training and validation loss
Training loss
Validation loss
2.5
2.0
1.5
1.0
0.5
0.0
0
20
40
60
80
100

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

A larger dataset consisting of 1500 images was utilized for instructing the computational model. Notwithstanding, a subset of 500 images was still employed for validation during the training phase, alongside an additional 500 images for testing purposes. Employing augmentation techniques such as image flipping, rotation, and zooming enhanced the program's learning capabilities significantly. Consequently, the program exhibited improved performance following the implementation of these techniques. Throughout the training phase, it achieved an accuracy rate of approximately 50%, while during validation also it demonstrated an accuracy of around 50%.

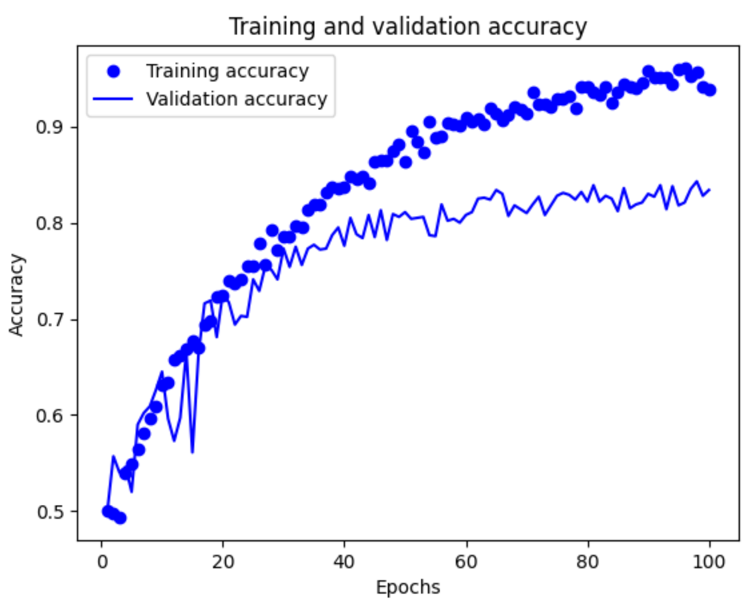

Training and validation accuracy
Training accuracy
Validation accuracy
Accuracy
Epochs
0.54
0.53
0.52
0.51
0.50
0.49
0.48
0.47
0
20
40
60
80
100

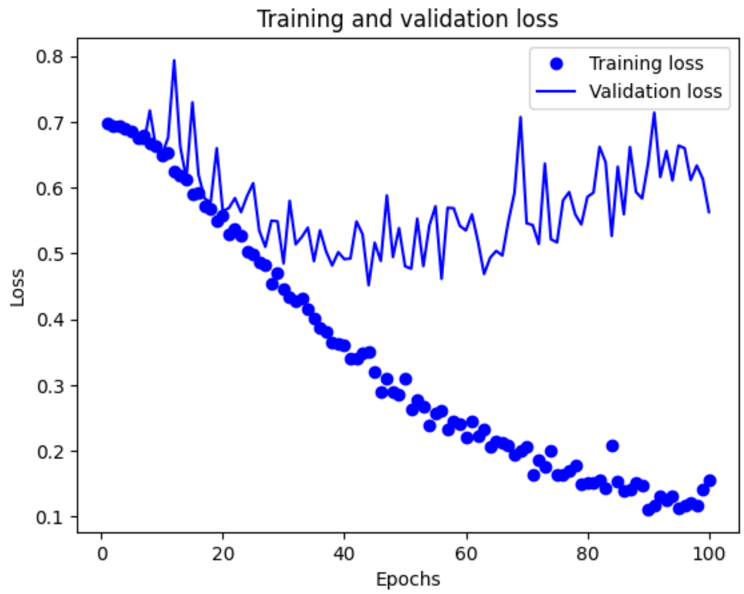

Training and validation loss
Training loss
Validation loss
Loss
Epochs
0.694
0.693
0.692
0.691
0.690
0.689
0
20
40
60
80
100

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 the previous steps. The objective is to find the ideal training sample size to get the best prediction results.

Answer:

A larger dataset comprising 2000 images was acquired to further enhance the computational model's performance. Augmentation techniques such as flipping, rotating, and zooming were consistently applied to these images during the training process. The integration of this expanded dataset and augmentation techniques notably improved the program's proficiency in image comprehension. Subsequently, during the training phase, the program exhibited an accuracy rate of approximately 93.8%, while during validation, it achieved an accuracy of around 83.4%.



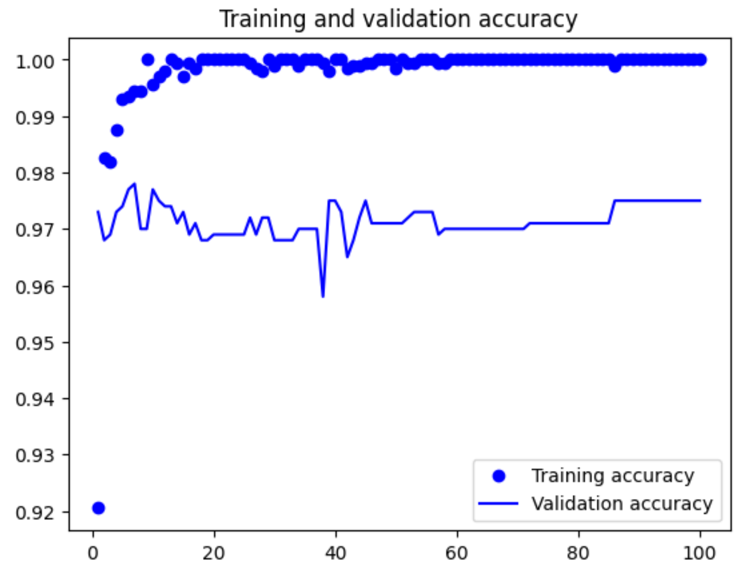


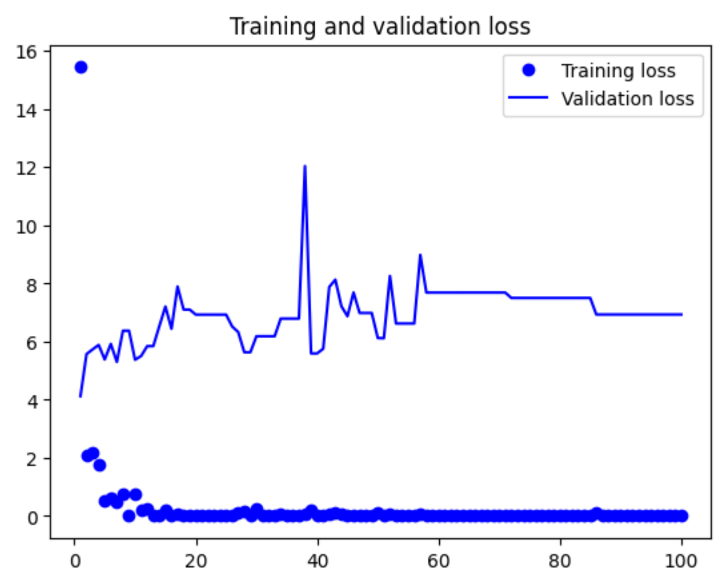
Q4: Repeat Steps 1-3, but now using a pre-trained network. The sample sizes you use in Steps 2 and 3 for the pre-trained network may be the same or different from those using the network where you trained from scratch. Again, use any and all optimization techniques to get the best performance.

Answer:

Pre-Trained Without Augmentation

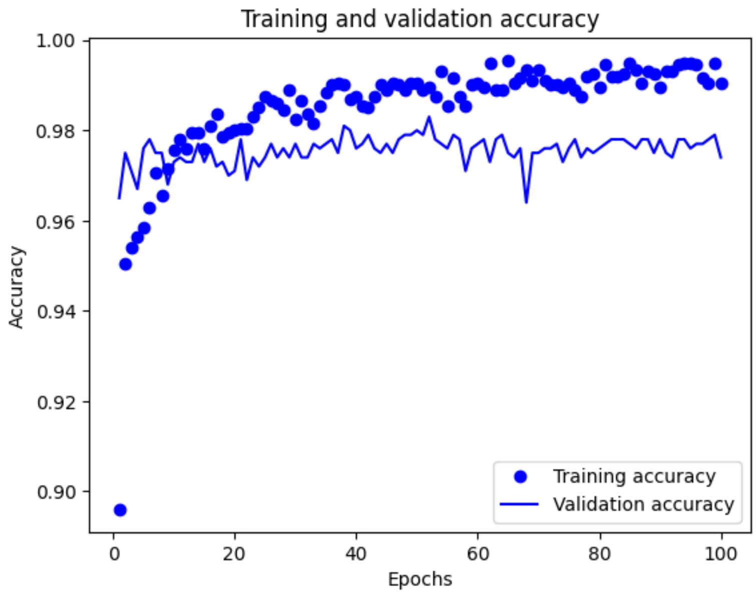
In the absence of augmentation techniques, we experimented with a pre-trained model for this inquiry, signifying the utilization of a model previously trained on a substantial volume of images. However, we refrained from applying any augmentation methods such as flipping or rotation to the images in this instance. Despite the absence of these techniques, the pre-trained model exhibited remarkable proficiency in image recognition. Throughout the training phase, it achieved a commendable accuracy rate of approximately 100%, which appears promising; however, this high accuracy might indicate an excessive reliance on the training dataset and insufficient adaptation to novel inputs. During validation, the model demonstrated an accuracy of about 97.5%, hinting at potential challenges in generalizing its performance beyond the training dataset.

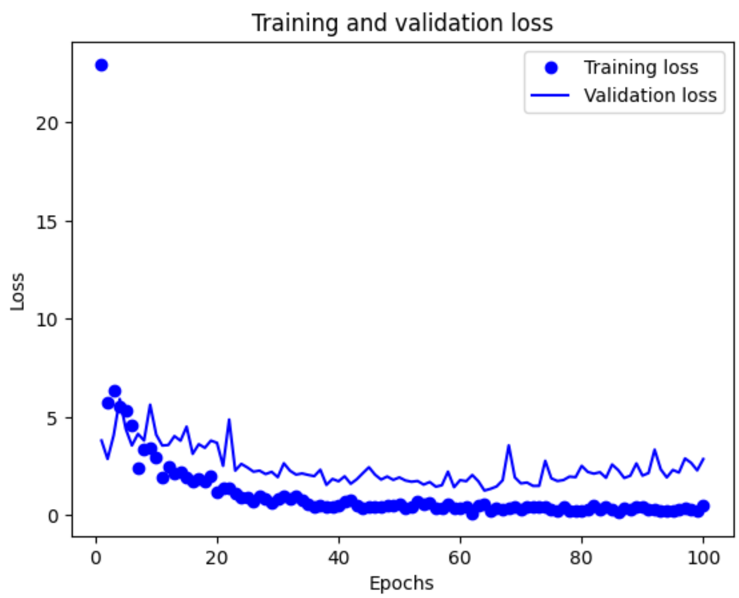




Pre-Trained With Augmentation:

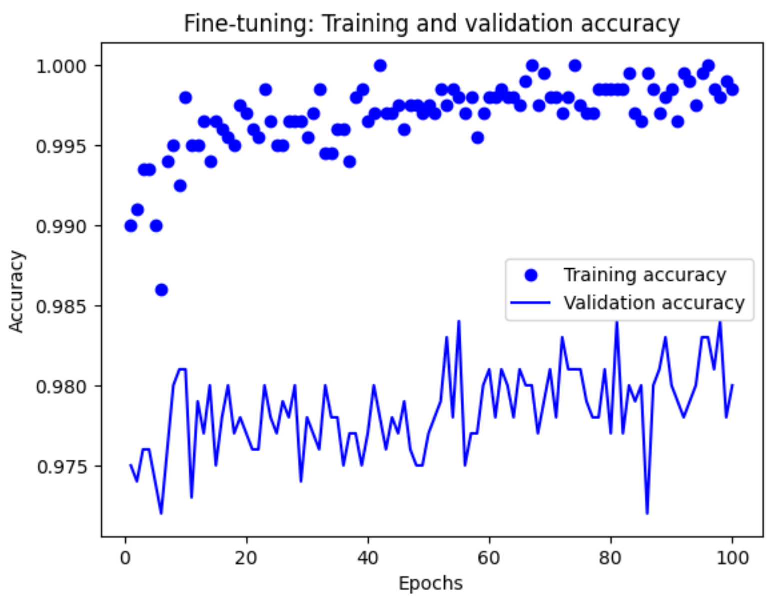
In the absence of employing any enhancements to augment the dataset, the pre-trained model exhibited commendable performance, attaining a validation accuracy of 97.4%. Subsequently, the author experimented with a technique known as fine-tuning, entailing minor adjustments to the pre-trained model to optimize its suitability for the specific task at hand. Following the fine-tuning process and the incorporation of data augmentation methods, the model displayed enhanced proficiency. Throughout the training phase, it achieved an accuracy rate of approximately 99.05%, while during validation, it demonstrated an accuracy of around 98%.

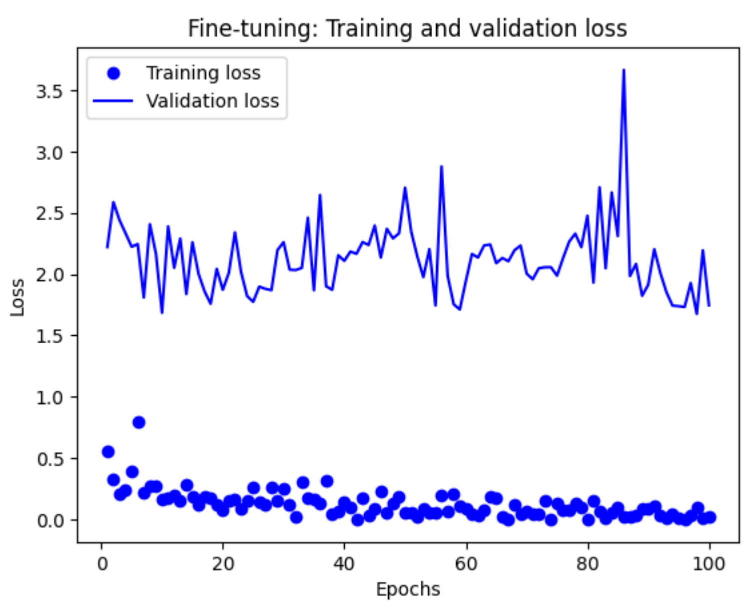




Fine-Tuning With Augmentation:

Following the experimentation with the pre-trained model alongside variations in the application of supplementary data enhancements, I proceeded to implement a technique known as fine-tuning to further refine its performance. Fine-tuning involves adjusting the pre-trained model to optimize its suitability for the particular task at hand. This process entailed allowing the layers of the pre-trained model to adapt to the newly enriched dataset, which had undergone augmentation techniques such as flipping and rotation. The fine-tuning procedure yielded significant improvements, resulting in heightened accuracy of the model. Throughout the training phase, it achieved an accuracy rate of approximately 99.8%, while during validation, it demonstrated an accuracy of around 98%. These findings underscore the efficacy of amalgamating a pre-trained model with data augmentation methods and fine-tuning, illustrating their collective impact on enhancing model performance. Notably, the fine-tuned model surpassed the performance of the pre-trained model without additional enhancements, underscoring the importance of diverse data and model customization tailored to the specific task at hand.





Conclusion:

The efficacy of the model is contingent upon the quality and quantity of the data it assimilates. Increasing the size of the training dataset from 1000 to 2000 images resulted in enhanced recognition capabilities, as evidenced by the rise in accuracy from 74.6% to 81% during testing. Furthermore, coupling a pre-trained model with techniques aimed at augmenting the dataset yields superior outcomes. Overall, the author posits that employing data enhancement techniques and expanding the dataset can facilitate more precise predictions and foster a deeper understanding of the subject matter by the model.