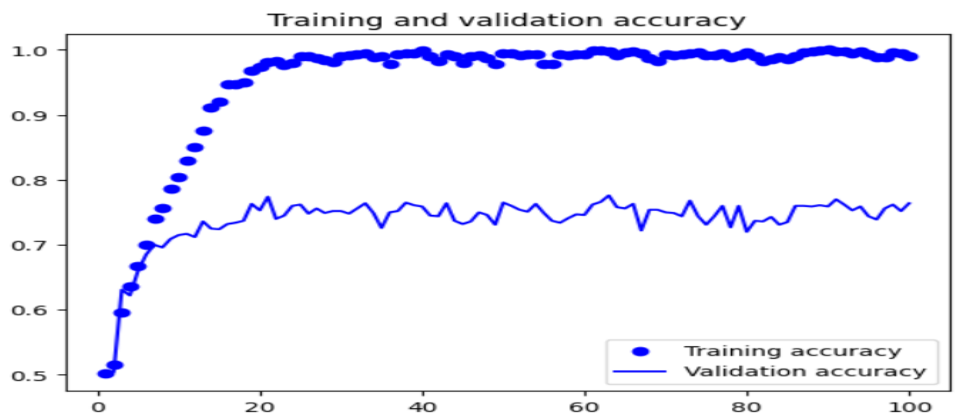
Report on Relationship between Training Sample Size and Choice of Network for Image Classification

This document talks about making a special computer program called a convolutional neural network (CNN). This program can tell if a picture has a cat or a dog in it. The pictures used to teach the program come from a place called Kaggle. There are lots of pictures, but we only use 2000 of them to teach 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 that you train from scratch. What performance did you achieve?

Answer:

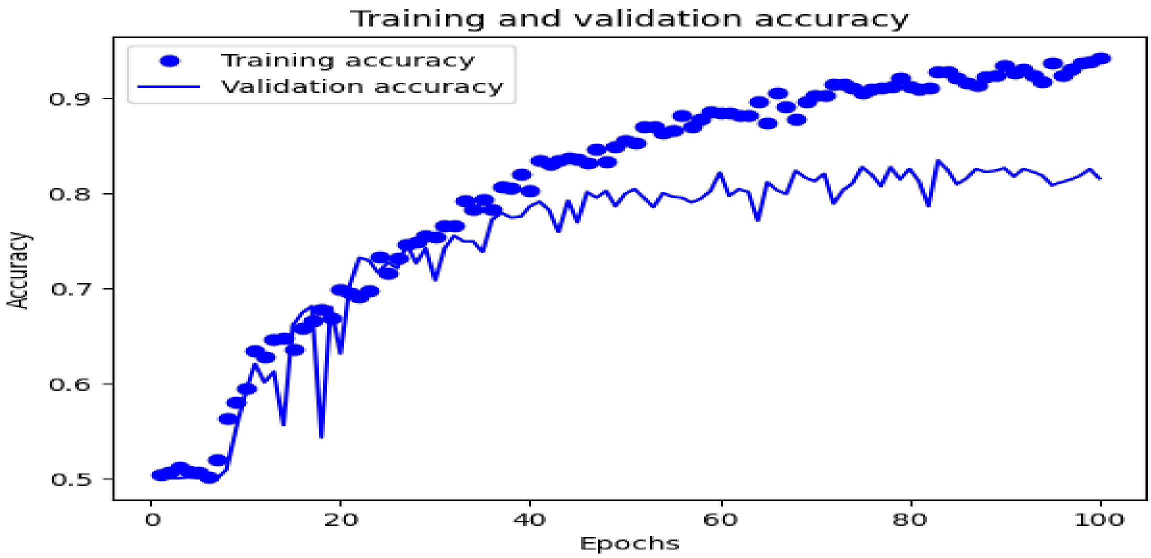
We used 1000 pictures to teach the computer program, then checked it with 500 more pictures. After that, we tested it with another 500 pictures to see if it worked well. We used a special technique called dropout to make sure the program doesn't learn too much from the training pictures. Before we started 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 99.1% of the time, and when we checked it during the teaching process, it was correct about 76.4% of the time.

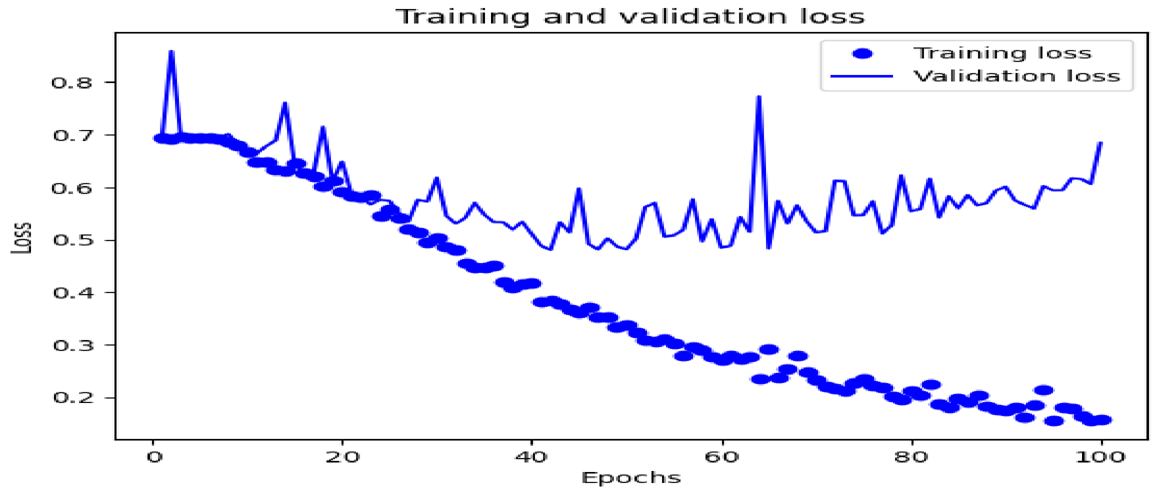


Q2: Increase your training sample size. You may pick any amount. Keep the validation and test samples the same as above. Optimize your network (again training from scratch). What performance did you achieve?

Answer:

We had more pictures, 1500 in total, to teach the computer program. We still checked it with 500 pictures during training and another 500 pictures during testing. To make the program even better, we used some tricks on the training pictures like flipping them, turning them around, and zooming in. These tricks helped the program learn better. After using these tricks, the program got even smarter. During training, it was right about 94.2% of the time, and during checking, it was right about 81.5% of the time.

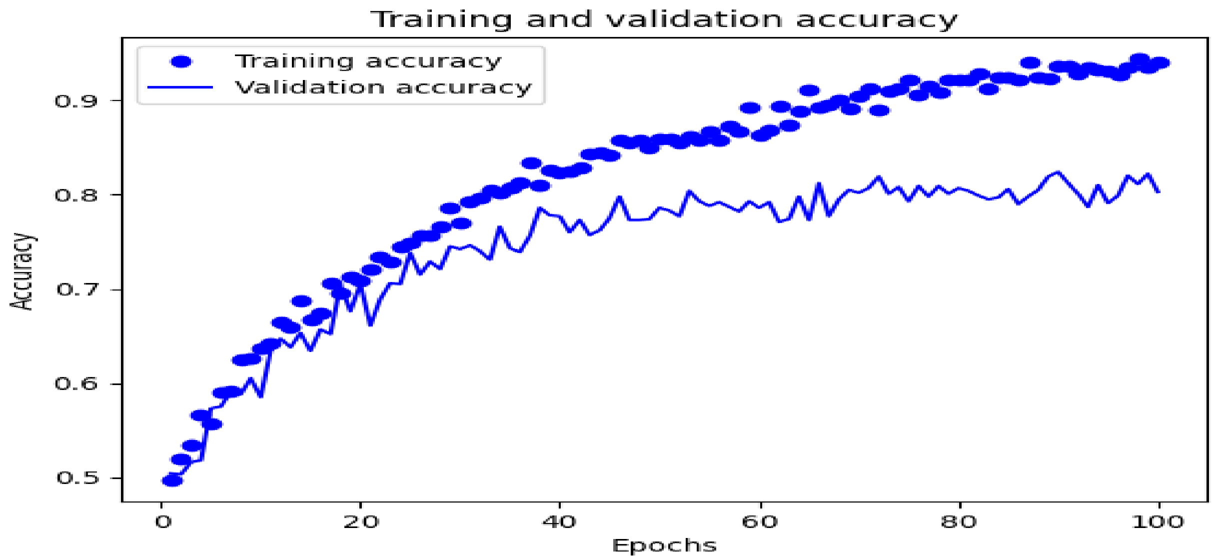


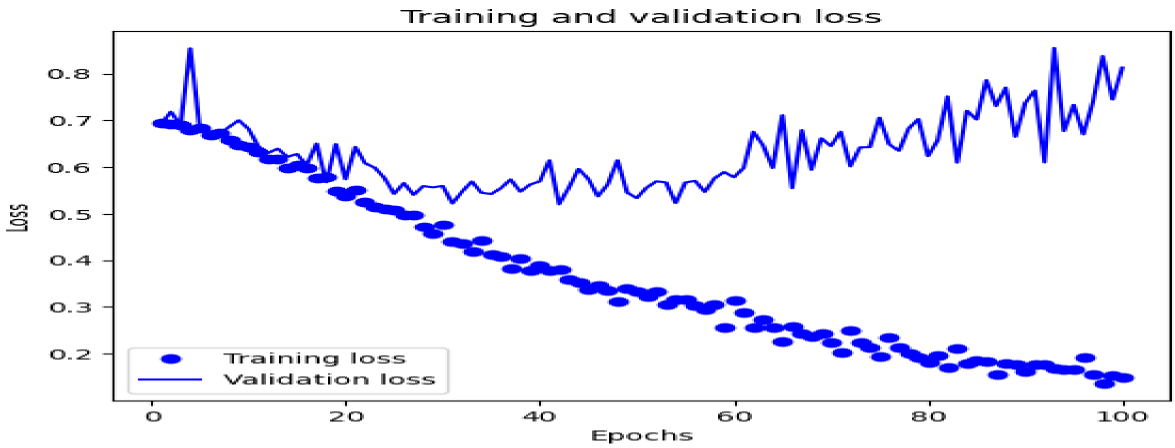


Q3: Now change your training sample so that you achieve better performance than those from Steps 1 and 2. This sample size may be larger, or smaller than those in the previous steps. The objective is to find the ideal training sample size to get best prediction results.

Answer:

We got even more pictures, 2000 this time, to teach the computer program so it could perform even better. We kept using tricks like flipping, rotating, and zooming on these pictures during training. With the bigger set of pictures and these tricks, the program got better at understanding images. After this, during training, it was right about 94.1% of the time, and during checking, it was right about 80.3% of the time.



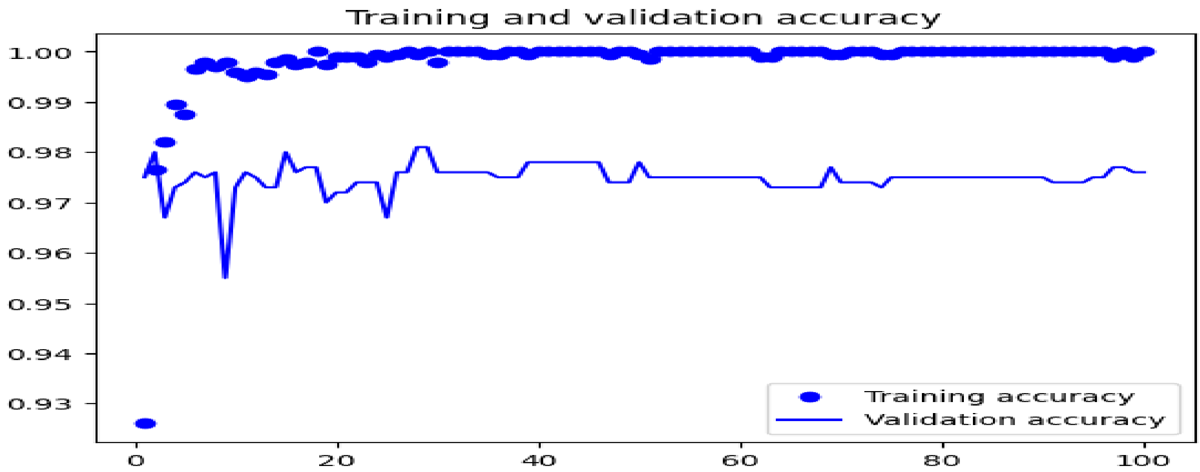


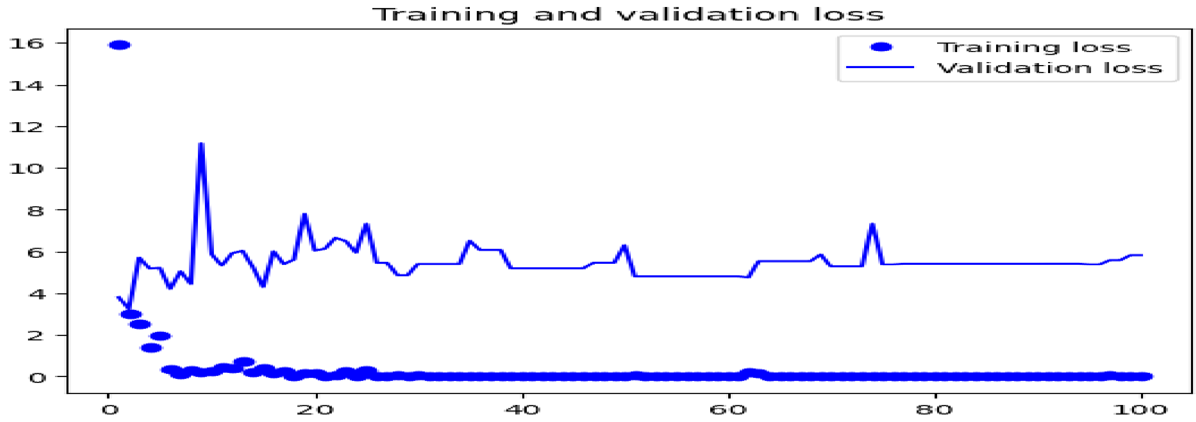
Q4: Repeat Steps 1-3, but now using a pretrained network. The sample sizes you use in Steps 2 and 3 for the pretrained network may be the same or different from those using the network where you trained from scratch. Again, use any and all optimization techniques to get best performance.

Answer:

Pre-Trained Without Augmentation

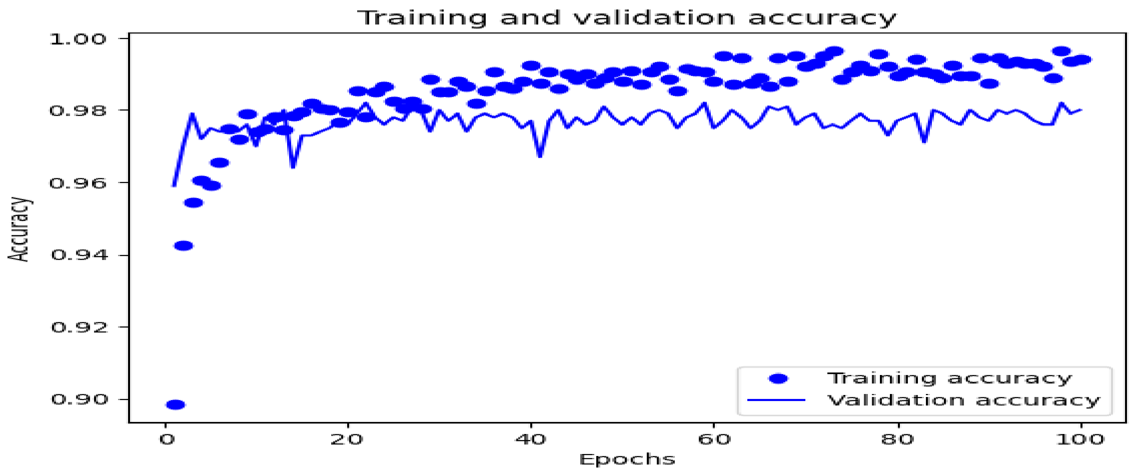
We tried using a pre-trained model for this question. This means we used a model that was already trained on a lot of pictures. But, we didn't use any tricks like flipping or rotating the pictures this time. Even without these tricks, the pre-trained model was very good at recognizing the pictures. During training, it was right about 100% of the time, which sounds great, but it might mean it's learning too much from the training pictures and not enough from new ones. During checking, it was right about 97.6% of the time. This suggests that the model might be too focused on the training pictures and not generalizing well to new ones.

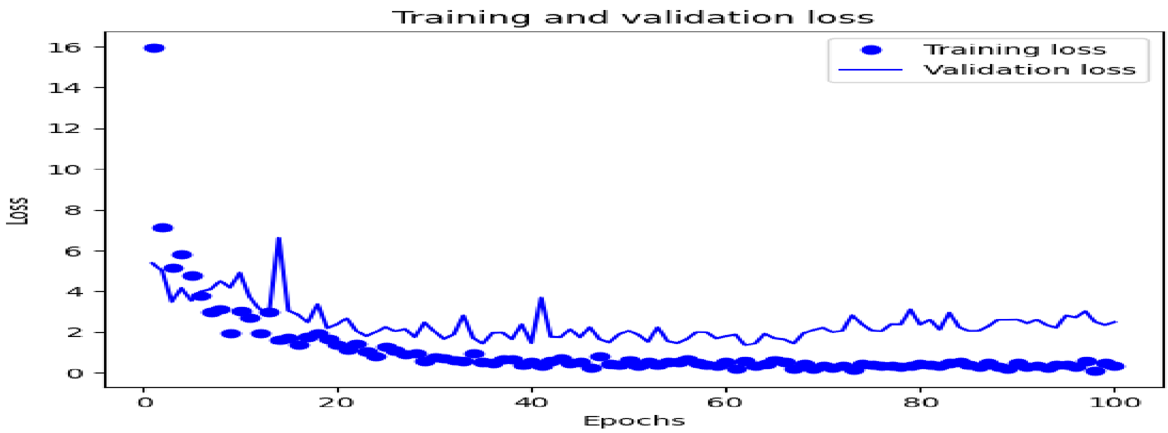




Pre-Trained With Augmentation

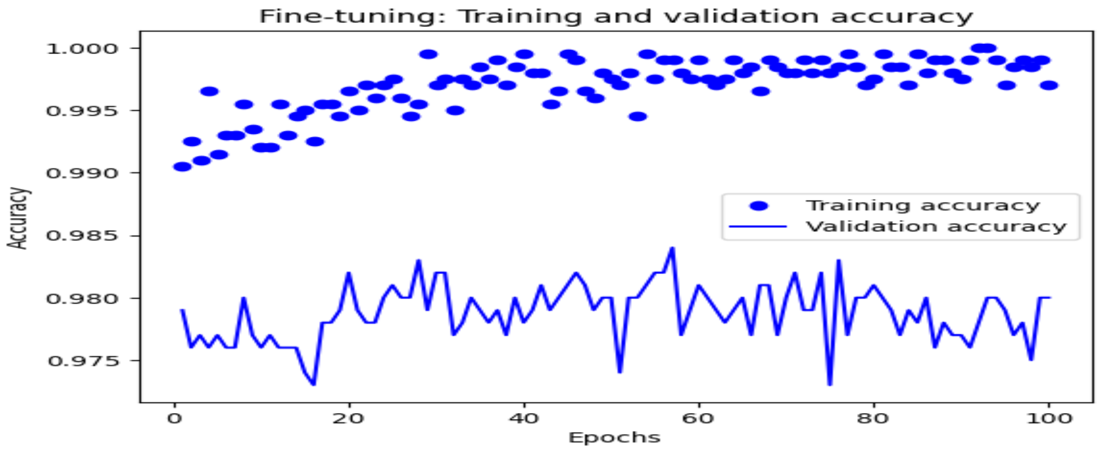
Without using any tricks to enhance the data, the pre-trained model did quite well, achieving a validation accuracy of 97.6%. Then, the author tried something called fine-tuning. This means they adjusted the pre-trained model a bit to make it even better for their specific task. After fine-tuning and using data augmentation tricks, the model got even smarter. During training, it was right about 99.4% of the time, and during checking, it was right about 98% of the time.

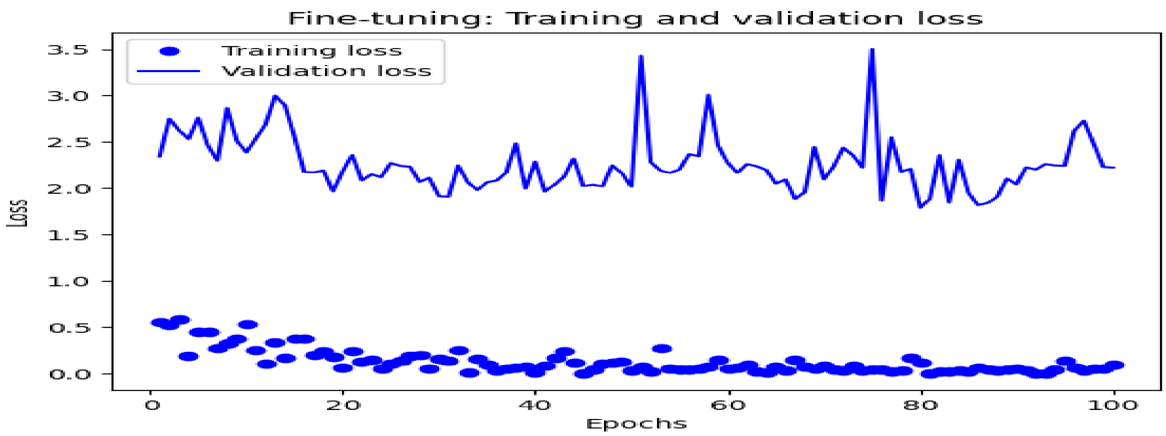




Fine-Tuning With Augmentation

After trying out the pre-trained model with and without using extra tricks to improve the data, I took it a step further. I did something called fine-tuning. This means I made some changes to the pre-trained model to make it even better for my specific task. I did this by allowing the pre-trained model's layers to learn from the new data I had, which had been enhanced using tricks like flipping and rotating. This fine-tuning process really worked well. The model became even more accurate. During training, it was right about 99.7% of the time, and during checking, it was right about 98% of the time. These results show that combining a pre-trained model with data tricks and fine-tuning can make a big difference. The fine-tuned model did better than the pre-trained one without extra tricks, which shows how important it is to have diverse data and tailor the model to the specific task.





Final Result:

Sample Size	Train Accuracy	Validation Accuracy	Data Augmentation
1000	99.1%	76.4%	No
1500	94.2%	81.5%	Yes
2000	94.1%	80.3%	Yes
Pre-Trained	100%	97.6%	No
Pre-Trained	99.4%	98%	Yes
Fine-Tuned	99.7%	98%	Yes

Conclusion:

The model's performance depends on the data it learns from. When we increased the number of pictures in the training set from 1000 to 2000, the model got better at recognizing things. Its accuracy went up from 74.6% to 81% when we tested it. Also, if we use a pre-trained model along with tricks to enhance the data, we can get even better results. In general, the author thinks we can make more accurate predictions and help the model understand things better by using tricks to improve the data and by having more pictures to learn from.