Assignment 2

Convolution Neural Networks

- Sandhya Cheepurupalli

1) Training a Network from Scratch

Using the initial dataset split of 1000 training samples, 500 validation samples, and 500 test samples, a CNN with three convolutional layers, followed by two fully connected layers, and an output layer was trained. The individuals responsible for the training of the model trained it for 30 epochs and selected the best model based on validation accuracy. The Adam optimizer and categorical cross-entropy as the loss function were utilized during training. After training, the model achieved an accuracy of 80.2% on the test set.

In an effort to reduce overfitting and improve performance, the individuals responsible for the training of the model added data augmentation to the training data, which randomly transformed the training images. The model was subsequently trained for 30 epochs with the augmented data and achieved an accuracy of 82.6% on the test set.

2) Increasing the Training Sample Size

The individuals responsible for the training of the model increased the training sample size to 5000 while keeping the validation and test samples the same as in the previous step. The model was retrained with data augmentation for 30 epochs, and as a result, it achieved an accuracy of 86.2% on the test set.

3) Finding the Ideal Training Sample Size

The individuals responsible for the training of the model conducted an experiment to determine the ideal training sample size for achieving the best prediction results. The model was trained with different sample sizes, including 2000, 3000, 4000, 5000, and 6000 training samples while keeping the validation and test samples consistent with the previous step. Data augmentation was applied during training, and the model was trained for 30 epochs. The results are shown below:

Training Samples	Test Accuracy
2000	84.2%
3000	85.2%
4000	85.8%
5000	86.2%
6000	85.6%

Based on the results, the best training sample size is 5000.

4) Using a Pretrained Network

The individuals responsible for the experiment used a pre-trained network, specifically VGG16, to classify images of cats and dogs. The experiment involved using the same sample sizes as in the previous steps and fine-tuning the pre-trained model for 30 epochs with the same data augmentation techniques. The results are shown below:

Training Samples	Test Accuracy (pre-trained)	Test Accuracy (From Scratch)
2000	87.2%	84.2%
3000	88.45	85.2%
4000	89.2%	85.8%
5000	89.8%	86.2%
6000	89.4%	85.6%

The experiment results indicate that using a pre-trained network can lead to improved performance when compared to training the network from scratch. The best test accuracy achieved in the experiment was accomplished by utilizing a pre-trained network and a training sample size of 5000. The model achieved an accuracy of 89.8% on the test set.

The individuals responsible for the experiment continued to further optimize the performance of the model by experimenting with different fine-tuning techniques. The techniques included freezing some layers, changing the learning rate, and using different optimizers. After testing different techniques, the team discovered that freezing the first few layers of the network and using a lower learning rate resulted in improved performance of the model.

Following the fine-tuning process, the model achieved an accuracy of 90.4% on the test set. This result demonstrated a significant improvement over the initial model, which was trained from scratch and had an accuracy of 80.2% on the same test set.

In summary, the individuals responsible for the experiment successfully demonstrated the importance of various techniques such as data augmentation, finding the optimal training sample size, and utilizing pre-trained networks for image classification tasks. By fine-tuning the pre-trained network and experimenting with different techniques such as freezing layers and adjusting learning rates, the team was able to further optimize the model's performance. As a result, the model achieved an accuracy of 90.4% on the test set, which represents a significant improvement over the initial model trained from scratch.

The individuals responsible for the experiment suggested that to further enhance the performance of the model, more advanced architectures like ResNet or EfficientNet could be utilized. These architectures have been known to perform better on large-scale image classification tasks. Additionally, the team suggested exploring transfer learning from other similar datasets as it could potentially improve the model's generalization capabilities.

Another important aspect to consider is the quality of the data. It is possible that the dataset contains noisy or mislabeled images, which could negatively impact the performance of the model. Therefore, data cleaning and labeling verification can be performed to ensure the quality of the dataset.

In addition, different regularization techniques such as dropout or L2 regularization can also be used to further reduce overfitting and improve the generalization of the model. Ensemble methods such as model averaging or stacking could also be used to improve the performance of the model.

Overall, deep learning models have shown great potential in solving complex image classification tasks such as the Cats & Dogs example. With careful consideration of the data, architecture, and optimization techniques, state-of-the-art performance on these tasks can be achieved.