Assignment 3

Convolution Neural Networks

1)Training a Network from Scratch

The original dataset was divided into 1000 training samples, 500 validation samples, and 500 test samples. A CNN consisting of three convolutional layers, two fully connected layers, and an output layer was trained using this split. The model was trained for 30 epochs by a team of individuals who selected the best model based on validation accuracy.

The training process utilized the Adam optimizer and categorical cross-entropy as the loss function. After the training process, the model achieved an accuracy of 80.2% on the test set. To address overfitting and enhance performance, the training data was augmented with random transformations of the images. The model was then trained again 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 individual in charge of training the model enlarged the training sample to 5000 but left the validation and test samples unchanged from the previous step. After retraining the model with data augmentation for 30 epochs, it was able to achieve a test set accuracy of 86.2%.

3)Finding the Ideal Training Sample Size

The people in charge of training the model performed an investigation to determine the optimum training sample size for achieving the most accurate predictions. They trained the model using various sample sizes, such as 2000, 3000, 4000, 5000, and 6000, while maintaining the same validation and test samples as in the previous stage. They also applied data augmentation during training and trained the model for 30 epochs. The following outcomes were observed:

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 people conducting the experiment employed a pre-existing neural network, VGG16, to categorize images of cats and dogs. They utilized the same sample sizes as in the previous steps and fine-tuned the pre-trained model by applying the same data augmentation methods for 30 epochs. The outcomes are presented 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%

According to the results of the experiment, it was observed that utilizing a pre-trained network can result in better performance as compared to training the network from the beginning. The highest accuracy achieved in the experiment was obtained by using a pre-trained network along with a training sample size of 5000. The model's accuracy on the test set was 89.8%.

The team conducting the experiment continued to refine the model's performance by exploring various fine-tuning methods, such as modifying the learning rate, using different optimizers, and freezing specific layers. Upon testing different approaches, the team found that the best results were achieved by freezing the initial layers of the network and using a lower learning rate.

After the fine-tuning process, the model was able to reach an accuracy of 90.4% on the test set. This outcome showed notable progress in comparison to the initial model that was trained from scratch, which only achieved an accuracy of 80.2% on the same test set.

To sum up, the experiment conducted by the team successfully highlighted the significance of various techniques, including data augmentation, finding the right sample size, and using pre-trained networks for image classification tasks. Additionally, by fine-tuning the pre-trained network and testing different techniques like layer freezing and learning rate adjustments, the team was able to further optimize the model's performance. As a result, the final model achieved an accuracy of 90.4% on the test set, which is a noteworthy improvement compared to the initial model that was trained from scratch.

The experimenters proposed that to improve the model's performance further, more advanced architectures such as Resent or Efficient Net could be used. These architectures are known to perform better on large-scale image classification tasks. Moreover, they suggested exploring transfer learning from other similar datasets as it could potentially enhance the model's ability to generalize.

The standard of the data is a further crucial factor to consider. The dataset might include noisy or incorrectly labeled images, which would be detrimental to the model's performance. Therefore, the dataset's quality can be ensured by performing data cleaning and labeling verification.

To further reduce overfitting and enhance the generalization of the model, various regularization techniques can also be used, such as dropout or L2 regularization. The model's performance could also be enhanced by ensemble techniques like model averaging or stacking.

The ability of deep learning models to handle challenging image classification tasks like the Cats & Dogs example has been demonstrated overall. Modern performance on these tasks can be attained with careful consideration of the data, architecture, and optimization techniques.