Assignment 2: Convolution

KENT ID: 811292363

Using a portion of the popular "Dogs-vs-Cats" dataset on Kaggle gives me a challenging chance to create a very successful model with a small amount of data. Convolutional neural network, also referred to as convnets, are the preferred method in computer vision for image identification, object detection, and segmentation tasks because of its remarkable capacity to learn and identify spatial patterns in images. Even with the restricted amount of data provided, I am confident that I can get remarkable outcomes by employing convolutional neural networks' capacity to extract and recognize crucial features from images. My ideal workflow would be to use a small dataset for training, advanced transfer learning methods for fine-tuning, and suitable assessment criteria for assessing the model's performance. My goal is to create a convolutional neural network that is both accurate and efficient, capable of accurately classifying images from the "Dogs-vs-Cats" dataset with minimal input.

In this study, I aimed to compare the performance of convolutional neural networks (CNNs) trained from scratch versus using a pretrained model for classifying images from the "Dogs-vs-Cats" dataset. I have also explored the impact of data augmentation on model performance.

Dataset and Methodology:

The dataset comprised 25,000 images of dogs and cats, with 12,500 images in each class. I created a new dataset with 1000 samples for each class for training, 500 samples for each class for validation, and 500 samples for each class for testing. The images were resized to 150x150 pixels and converted to RGB pixel grids.

We used TensorFlow and Keras for model development and evaluation.

Training from Scratch:

I trained several models from scratch with varying sample sizes and data augmentation techniques.

Model 1 with 1000 samples per class achieved a test accuracy of 76.8% and a validation accuracy of 70.6% without data augmentation.

Model 2 with 1500 samples per class achieved a test accuracy of 83% and a validation accuracy of 71.9% without data augmentation.

Using Pretrained Models:

I have utilized the VGG16 pretrained model for transfer learning.

The pretrained model achieved a test accuracy of 99.6% and a validation accuracy of 97% without data augmentation.

With data augmentation, the pretrained model achieved a test accuracy of 95.8% and a validation accuracy of 97.2%.

Pre-trained model: -

The original dataset is huge and diverse, and its properties make it useful for a variety of computer vision applications. A pretrained network can be utilized as a generic model. The capacity of deep learning to transfer learned properties across tasks is one of its main advantages over other machine learning techniques.
Imagine a massive convolutional neural network that was trained on the 1,000 unique classes and 1.4 million annotated images in the ImageNet dataset. Numerous animal categories, including various cat and dog breeds, are included in this dataset. The architecture of this network, referred to as VGG16, is a fundamental and widely utilized convnet architecture for ImageNet.

Data Augmentation:

I suggest using data augmentation techniques to raise our model's accuracy. Even with limited datasets, we can obtain good results by randomly modifying the provided training samples to create new data. Consequently, this helps with generalization since the model will never see the same image twice while it is being trained.
To achieve our specific goal, I wish to randomly apply effects like flipping, rotating, and zooming to the training set of photographs. By doing this, we might produce variations of the existing images, adding diversity to the dataset and enhancing our model's resilience.

Technique:

This is a binary classification problem using the Cats-vs-Dogs dataset, where you have to guess if a photo is in the dog or cat class.

- Start the picture files.
- Transform the JPEG content into RGB pixel grids.
- Transform them into floating-point tensors.

The pixel values (ranging from 0 to 255) should be rescaled to the [0, 1] interval since, as you may know, neural networks prefer small input values.

The 543MB Cats-vs-Dogs dataset includes 25,000 images of dogs and cats, with 12,500 images in each class. (condensed). We'll create a new dataset with three subsets after downloading and unzipping it: a training set with 1000 samples for each class, a validation set with 500 samples for each class, and a test set with 500 samples for each class. Because the problem we're working on is more complex and has a larger image, we need to increase the neural network's capability. We plan to accomplish this by including a new stage into our current Conv2D + MaxPooling2D design. This will ensure that the feature maps are not overly large when we reach the Flatten layer by increasing network bandwidth while concurrently decreasing their size. The original dimensions of our input photographs are 150x150. As we move through the network layers, the feature maps get smaller and smaller until they are 7x7 right before the Flatten layer. Although the input size selection is rather arbitrary, it makes sense in this specific scenario.

Results for Model from Scratch

Model no	Train Size	Validation and Test sample size	Data Augmentation	Test Accuracy%	Validation Accuracy%
Model 1	1000	500,500	NO	76.8	70.6
Model 1a	1000	500,500	YES	67.1	64.2
Model 2	1500	500,500	NO	83	71.9
Model 2a	1500	500,500	YES	70.37	70.3
Model 2b	1500	500,500	YES	81.7	73.2
Model 2c	1500	500,500	NO	72.7	73.8

Results for Pre-Trained Models

Data Augmentation	Train Accuracy %	Validation Accuracy%
NO	99.6	97
YES	95.8	97.2

Conclusion:

The tables above contain the model settings and the sample sizes for the train, test, and validation sets. We present findings for the model trained from scratch, with and without data augmentation, and for models trained with different train and validation sizes. For the pre-trained model, we compare the accuracy, validation accuracy, and data augmentation.

The findings show that models trained with and without data augmentation did not consistently perform better than those trained without it. The accuracy of the model is also improved by expanding the training set or changing the validation set's size. The results of our comparison between the pre-trained model with and without data augmentation indicate that the accuracy of the model or its validation was not improved by the addition of data. In general, pre-trained models perform better than models built from scratch, particularly in situations when there is a dearth of training data.