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 "Dogsvs-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: -

Because of its size, diversity, and other qualities, the original dataset can be applied to a wide range of computer vision applications. A generic model could be a pretrained network. One of deep learning's primary advantages over other machine learning methods is its ability to transfer learnt attributes across tasks.
Consider a large-scale convolutional neural network trained on the 1.4 million annotated images and 1,000 distinct classes found in the ImageNet dataset. This dataset contains many animal classifications, such as different breeds of dogs and cats. This network's architecture, known as VGG16, is a basic and often used convnet architecture for ImageNet.

Data Augmentation:

I propose applying data augmentation approaches to improve the accuracy of our model. We can still get good results with small datasets by creating new data by randomly changing the available training examples. Because the model will never see the same image twice during training, this aids in generalization.
I want to sporadically add effects on the training set of photos, such as flipping, rotating, and zooming, in order to accomplish our particular aim. This could result in different versions of the current photos, bringing diversity to the dataset and strengthening the robustness of our model.

Technique:

With the help of the Cats-vs-Dogs dataset, you must determine whether a given photo belongs in the dog or cat class in this binary classification task.

Open the image files -> Convert the JPEG images into grids of RGB pixels -> Convert them to tensors with floating points.
It is necessary to rescale the pixel values (which range from 0 to 255) to the [0, 1] interval since neural networks, as you may be aware, favor small input values.

There are 25,000 photos of dogs and cats in the 543MB Cats-vs-Dogs dataset, with 12,500 photos in each class. (shortened). After downloading and unzipping the file, we'll construct a new dataset with three subsets: a test set with 500 samples per class, a validation set with 500 samples per class, and a training set with 1000 samples per class. The topic we are working on requires us to boost the neural network's capability because it is more complex and has a larger image. We intend to include a new stage into our existing Conv2D + MaxPooling2D system in order to achieve this. By boosting network bandwidth and decreasing feature map size at the same time, this will guarantee that the feature maps are not excessively huge when we get to the Flatten layer. The supplied photos have original dimensions of 150 x 150. The feature maps get smaller and smaller as we proceed through the network levels, reaching a size of 7 by 7 just prior to the Flatten layer. Even though the choice of input size is rather random, it makes sense in this particular situation.

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 pretrained 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.