CONVOLUTION  
ASSIGNMENT-2

Abhinav Reddy Yelipeddi

811300807

# Summary

This project's purpose is to design and build a new convolutional neural network specifically for computer vision applications. The dataset we are utilizing is based on the "Dog-vs-Cats" dataset available on Kaggle. Developing an effective model is difficult due to the restricted amount of data available. Convolutional neural networks, or convnets, are a common deep learning model that has proven to be highly effective in computer vision applications. Convnet's ability to discover and identify spatial patterns in images is one of its primary advantages. They excel in tasks such as segmentation, object identification, and image recognition.

We believe our convolution model can still produce good results despite the low amount of available data. This could be attributed to convolutional neural networks' ability to learn information and generalize from limited samples by extracting and recognizing significant characteristics from images. Our model will be trained on a limited dataset, fine-tuned with transfer learning methods, and its performance will be evaluated using appropriate criteria. Finally, we aim to build a convolutional neural network that can efficiently and accurately classify photos from the "Dog-vs-Cats" dataset with a small amount of input.

# Problem

# The objective of the Cats-vs-Dogs dataset binary classification task is to determine if an image is part of the dog or cat class.

# Techniques

## Dataset:

The Cats-vs-Dogs dataset is 543 MB in size (compressed) and includes 25,000 images of dogs and cats (12,500 from each class). Three subsets will be included in the new dataset we construct after downloading and unzipping it: a training set with 1000 samples of each class, a validation set with 500 samples of each class, and a test set with 500 samples of each class. We have to expand our neural network's size since the problem we are working on is more complicated and requires a larger picture. We will extend our current Conv2D + MaxPooling2D design by one stage to achieve this. This ensures that the feature maps are not too large when we get to the Flatten layer while simultaneously increasing the capacity of the network. First, our input images are 150x150 in size. The feature maps get smaller as we move up the network's layers, reaching 7x7 just before the Flatten layer. The input size selected is a bit arbitrary, but it fits the task at hand.

## Preprocessing:

* Read the picture files.
* Decode the JPEG content to RBG grids of pixels.
* Convert these into floating point tensors.
* Rescale the pixel values (between 0 and 255) to the [0, 1] interval (as you know, neural networks prefer to deal with small input values).

## Data Augmentation:

To improve the accuracy of our model, we propose to use data augmentation approaches. Data augmentation generates new data from training samples using random alterations, allowing us to achieve outstanding outcomes even with limited datasets. As a result, the model will never see the same image again while being trained, increasing its ability to generalize. To achieve our specific goal, we wish to conduct random operations on the photos in the training set, such as flipping, rotating, and zooming. This allows us to create different copies of previously existing photographs, increasing the variety of the dataset and the robustness of our model.

# Pre-trained model:

If the original dataset is broad and diverse, a pretrained network can be utilized as a generic model with features that can be used to a variety of computer vision applications. One of the primary advantages of deep learning over other machine learning methods is its ability to transfer learned properties between tasks. As an example, a huge convolutional neural network trained on the ImageNet dataset, which has 1.4 million annotated images and 1,000 unique classifications, can be evaluated. The collection contains many animal categories, including various dog and cat breeds. The architecture of this network is named VGG16, after a popular and basic convnet design for ImageNet.

Pretrained networks are mostly used for feature extraction and fine-tuning. To improve on the results, we will use feature extraction in this scenario, first without data augmentation and subsequently with data augmentation.

**Results:** The table below shows the accuracy and validation loss for each approach.

TABLE FOR MODEL FROM SCRATCH

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Train Size | Test Size | Validation Size | Data Augmentation | Train Accuracy(%) | Validation Accuracy  (%) |
| 1000 | 500 | 500 | NO | 77.4 | 71.6 |
| 1000 | 500 | 500 | YES | 70.3 | 65.6 |
| 1500 | 500 | 500 | NO | 84.5 | 75.4 |
| 1500 | 500 | 500 | YES | 71.5 | 71.2 |
| 1500 | 1000 | 500 | YES | 82.6 | 75.4 |
| 1500 | 1000 | 500 | NO | 67.4 | 66.2 |

TABLE FOR PRE-TRAINED MODEL

|  |  |  |
| --- | --- | --- |
| Data Augmentation | Train Accuracy(%) | Validation Accuracy(%) |
| NO | 99.7 | 97.8 |
| YES | 96.1 | 97.5 |

The sample sizes for the train, test, and validation sets are listed in the tables above, along with the model settings. We provide findings for the scratch-built model with and without data augmentation, as well as results for models trained with different train and validation sizes or with a larger train size. We compare the validation accuracy and accuracy of the pre-trained model with and without data augmentation.

Based on the findings, we can conclude that the models consistently trained with and without data augmentation did not outperform the others. The model's accuracy can also be improved by increasing the training set or changing the dimensions of the validation set. Data augmentation did not improve the pre-trained model's accuracy or validation accuracy, as evidenced by a comparison of the two sets of data. When working with minimal training data, pre-trained models frequently outperform models constructed from scratch.

**Conclusion:**

* This project successfully built and implemented a convolutional neural network for binary image classification using the "Cats-vs-Dogs" dataset.
* We investigated two basic approaches: building new models from scratch and fine-tuning pre-trained ones. The pre-trained VGG16 model performed well.
* Pre-trained models, particularly those using transfer learning, outperformed models constructed from scratch, especially when data availability was restricted.
* Data augmentation was tested to improve performance. However, it had little effect on the pre-trained model's accuracy, demonstrating that the advantage of augmentation varies depending on the model and dataset sizes.
* Increasing the training set size or altering the validation set proportions can help to enhance model accuracy and resilience.
* Future research could concentrate on sophisticated augmentation approaches and experiments with other architectures to improve model efficiency.

**THANK YOU !**