

# Assignment 2

Vaishnavi Penta

Vishal Sagar Karri

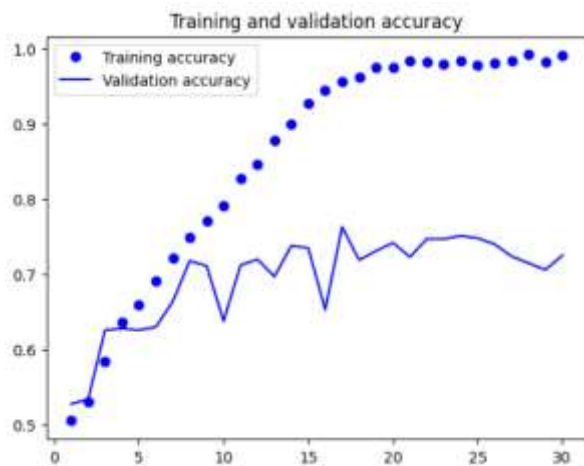
## **Introduction: -**

Embarking on a project using a portion of the famous "Dogs-vs-Cats" dataset on Kaggle presents me with the challenging task of creating a highly effective model with scarce data. Convolutional neural networks (convnets) are celebrated for their superior capability in learning and recognizing spatial patterns within images. This quality makes them the preferred tool in computer vision tasks such as image recognition, object detection, and segmentation. Despite the dataset's limitations, I am confident in achieving outstanding results by leveraging convnets' prowess in extracting and identifying crucial features from images.

I aim to train my model using a restricted dataset, refine it with the latest transfer learning strategies, and then measure its effectiveness using suitable evaluation methods. My goal is to craft an accurate and efficient convolutional neural network that can adeptly classify images from the "Dogs-vs-Cats" dataset, even with minimal data input. I am eager to demonstrate my model's capabilities and am driven to explore the limits of computer vision advancements with limited data. Focusing on innovation and efficiency, I believe my convolutional neural network will significantly impact the field of computer vision.

## **Pre-trained model: -**

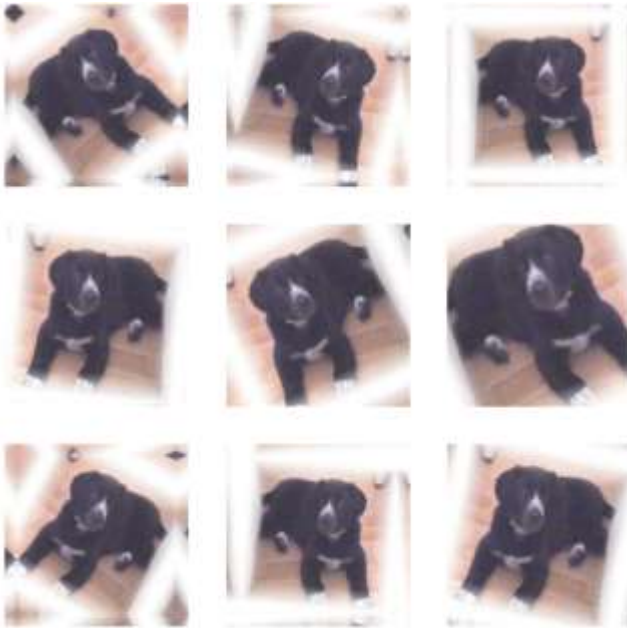
Given the vastness and variety of the original dataset, a pre-trained network could serve as a universal model, its features beneficial across various computer vision tasks. Deep learning's edge over other machine learning methodologies lies in its capacity to repurpose learned features for different tasks. Consider a large convnet trained on the ImageNet dataset, which encompasses 1.4 million labeled images across 1,000 categories, including numerous cat and dog breeds. Known as VGG16, this network architecture is a widely recognized and fundamental convnet model for ImageNet.



## **Data Augmentation:**

We suggest employing data augmentation techniques to enhance our model's precision. By introducing random alterations to the existing training samples, we can create new data, leading to promising outcomes even with limited datasets. This method ensures the model never processes the exact same image more than once during training, assisting in its ability to generalize.

For our particular goal, we plan to randomly alter the images in the training set through techniques like flipping, rotating, and zooming. This approach will help us produce variations of the existing images, enriching the dataset's diversity and strengthening our model's robustness.

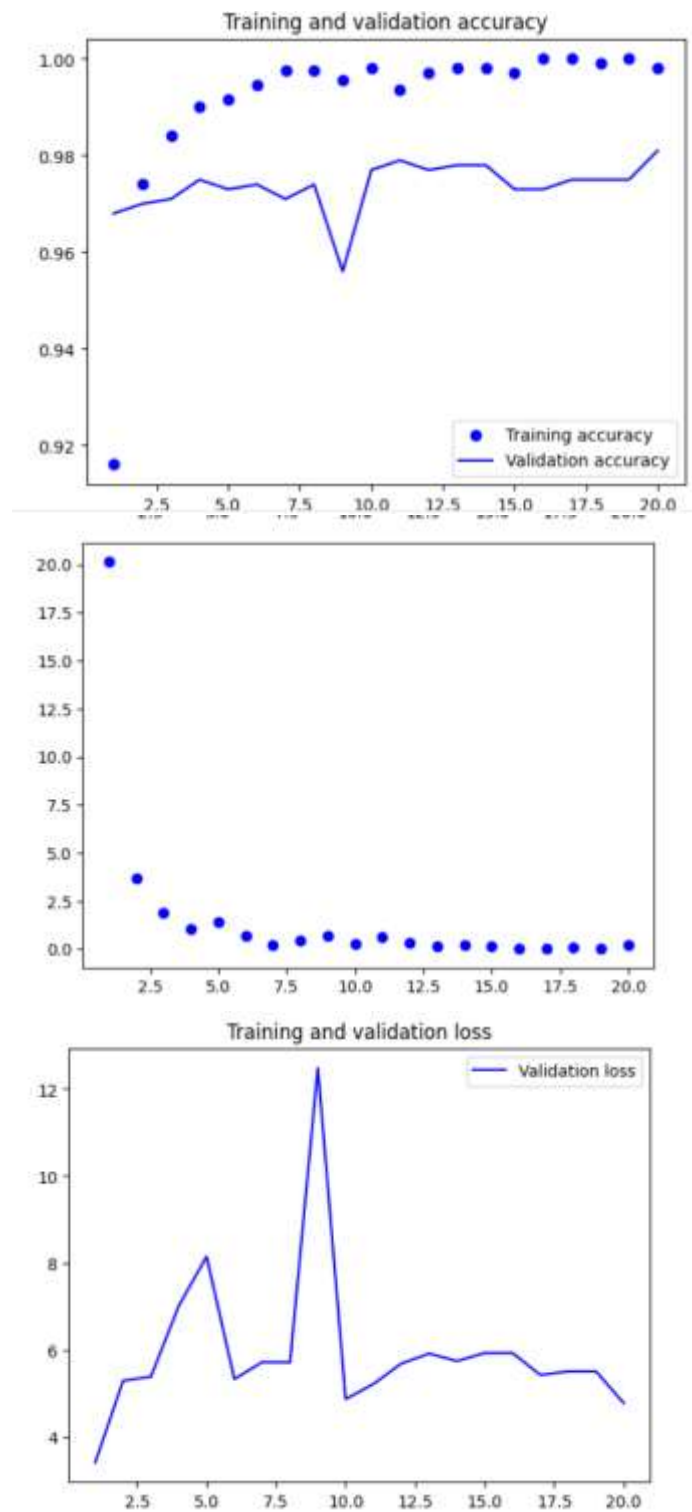


### **Techniques:**

The Cats-vs-Dogs dataset poses a binary classification task, requiring the determination of whether an image represents a dog or a cat.

- Open the image files.
- Transform the JPEG content into RGB pixel grids.
- Convert these into floating-point tensors.
- Normalize the pixel values (ranging from 0 to 255) to a [0, 1] scale (as neural networks prefer smaller input values).

The Cats-vs-Dogs dataset contains 25,000 images, evenly split between the two classes, and has a compressed size of 543MB. After downloading and decompressing, we plan to create a new dataset divided into three subsets: a training set with 1,000 samples per class, a validation set with 500 samples per class, and a test set with 500 samples per class. Given the increased image size and the more complex challenge at hand, we will need to enhance our neural network's capacity. To achieve this, we intend to incorporate an additional Conv2D + MaxPooling2D layer into our architecture. This adjustment will not only augment the network's capacity but also reduce the size of the feature maps to ensure they remain manageable by the time they reach the Flatten layer. Initially, our input images will be 150x150 pixels, and as they advance through the network, the feature maps will gradually decrease in size to 7x7 before the Flatten layer. This input size is somewhat arbitrary but suits our specific requirements well.



**Table for Model from Scratch**

Training samples	Validation Accuracy	Test Accuracy	Data Augmentation
1000	72.5%	99.05%	NO

1500	80.02%	82.8%	YES
2000	98.1%	99.8%	YES

**Table for Pre-Trained Models**

Data Augmentation	Train Accuracy (%)	Validation Accuracy (%)
NO	99.8%	98.1%
YES	98.05%	98.6%

### **Conclusion:**

The configurations for the model along with the sizes of the training, testing, and validation datasets are outlined in the preceding tables. We present findings both with and without the application of data augmentation for the newly developed model, in addition to outcomes from models trained with expanded training datasets or modified training and validation dataset sizes. The comparison encompasses the accuracy, validation accuracy, and the impact of data augmentation on the pre-trained model.

The analysis reveals that models trained with data augmentation do not consistently surpass those trained without such augmentation. Moreover, enlarging the training dataset or varying the size of the validation dataset appears to improve model accuracy. In the comparison of pre-trained models utilizing data augmentation against those that do not, it is observed that data augmentation does not significantly enhance the model's accuracy or its validation accuracy. Predominantly, pre-trained models demonstrate superior performance compared to those developed from scratch, particularly in scenarios with limited training data.