Assignment 2
Sai Prasad Boyapati

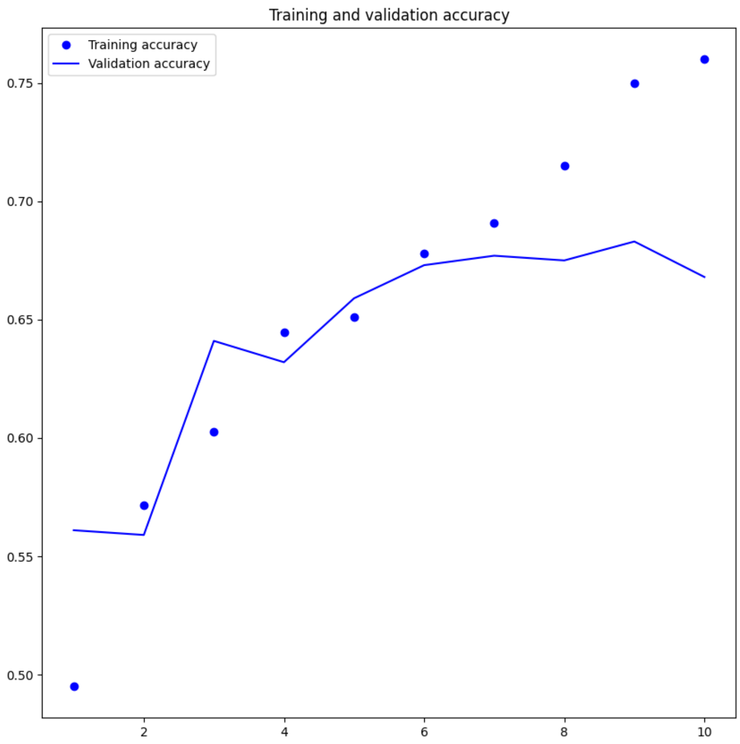
Introduction: -

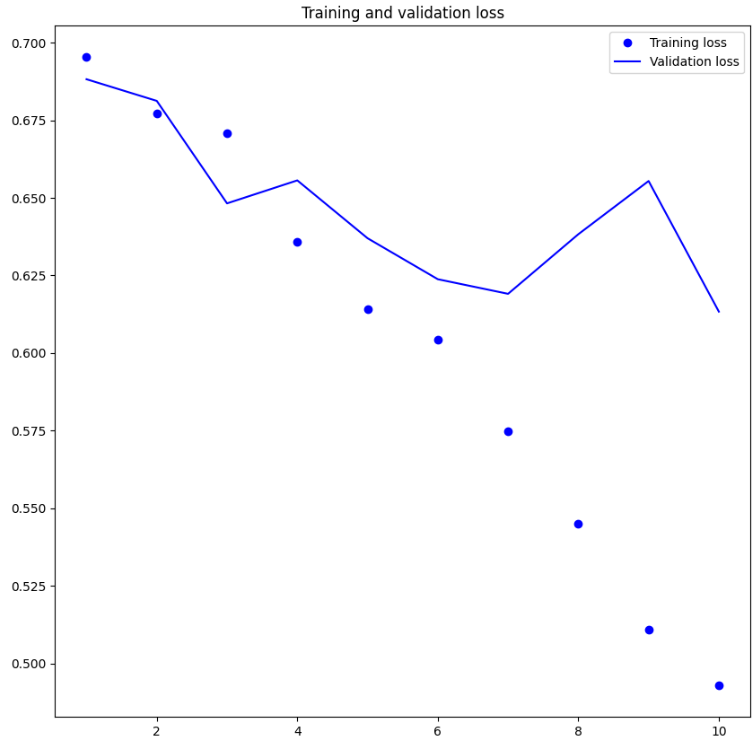
I'm starting a project with a portion of the well-known "Dogs-vs-Cats" dataset from Kaggle, aiming to build an effective model despite having limited data. Convolutional neural networks (CNNs) are renowned for their ability to learn and recognize spatial patterns in images, making them ideal for tasks like image recognition, object detection, and segmentation. I'm optimistic about achieving strong results by harnessing the strengths of CNNs to extract and identify key features from images.

My plan is to train the model using a small dataset, enhance it with the latest transfer learning techniques, and evaluate its performance with appropriate metrics. My objective is to develop a precise and efficient convolutional neural network capable of classifying images from the "Dogs-vs-Cats" dataset, even with minimal input data. I'm excited to showcase my model's capabilities and explore the potential of computer vision advancements with limited resources. By focusing on innovation and efficiency, I believe my CNN will make a meaningful contribution to the field of computer vision.

Pre-trained model: -

Given the extensive and diverse nature of the original dataset, using a pre-trained network can act as a versatile model, offering valuable features applicable to various computer vision tasks. The advantage of deep learning over traditional machine learning approaches is its ability to reuse learned features for different applications. For example, consider a large convolutional neural network trained on the ImageNet dataset, which includes 1.4 million labeled images across 1,000 categories, featuring many breeds of cats and dogs. VGG16, a well-known and foundational architecture for ImageNet, exemplifies this approach.





Data Augmentation:

We recommend using data augmentation techniques to improve our model's accuracy. By applying random modifications to the existing training samples, we can generate new data, which can yield promising results even with limited datasets. This approach ensures that the model does not encounter the same image more than once during training, aiding its ability to generalize.

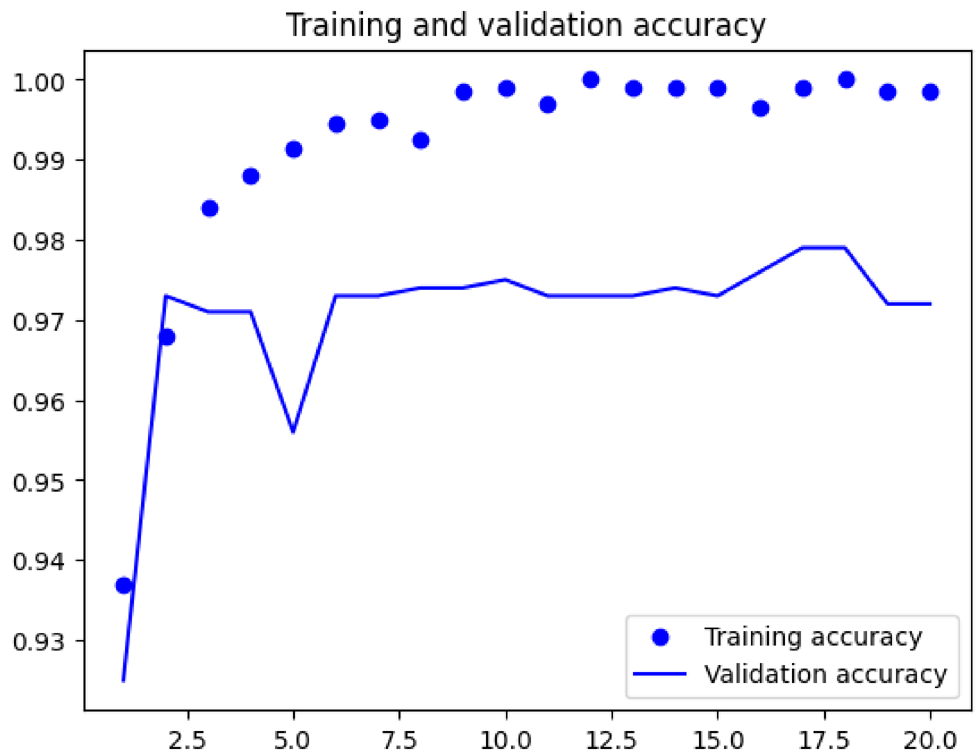
For our specific objectives, we plan to randomly modify the images in the training set using techniques such as flipping, rotating, and zooming. This will help create variations of the existing images, enhancing the diversity of the dataset and strengthening the robustness of our model.

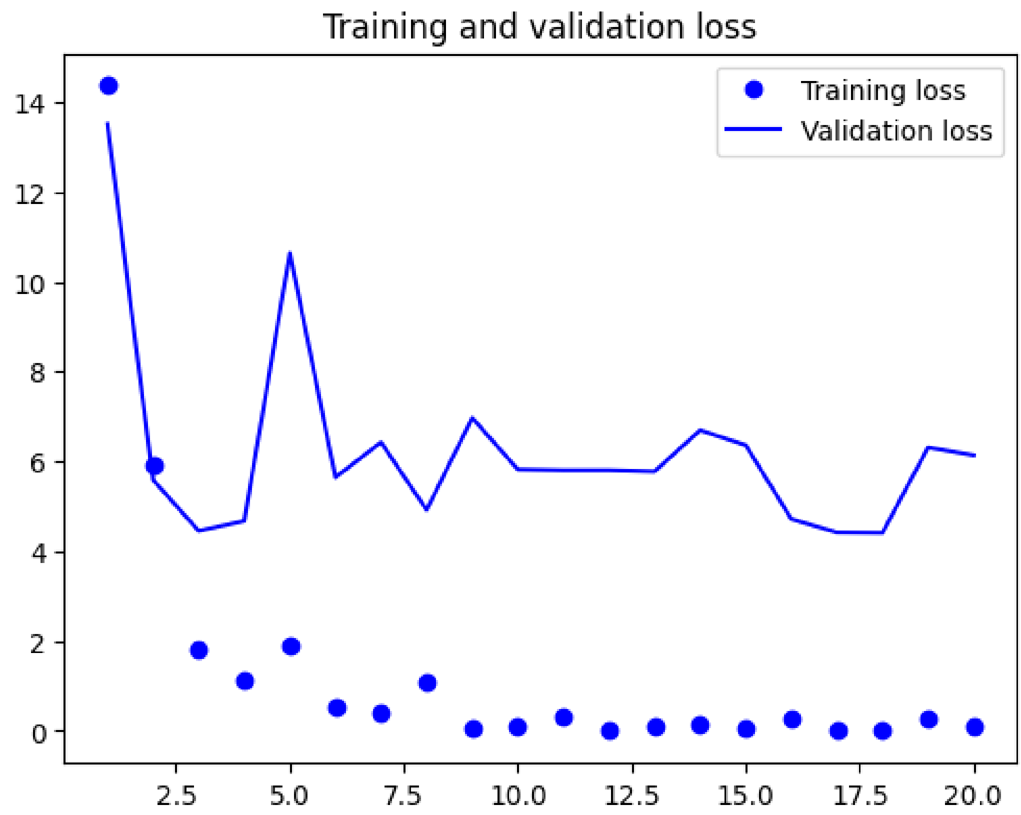
Techniques: The Cats-vs-Dogs dataset presents a binary classification challenge, requiring us to determine whether an image depicts a dog or a cat. Our process will involve opening the image files, transforming the JPEG content into RGB pixel grids, and converting these into floating-point tensors. We will normalize the pixel values (which range from 0 to 255) to a [0, 1] scale, as neural networks perform better with smaller input values.

The Cats-vs-Dogs dataset consists of 25,000 images, evenly divided between the two classes, and has a compressed size of 543MB. After downloading and decompressing, we plan to create a new dataset split 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 larger image size and the increased complexity of the task, we will need to enhance the capacity of our neural network. To achieve this, we intend to add an additional Conv2D + MaxPooling2D layer to our architecture. This modification will increase the network's capacity while also reducing the size of the feature maps, ensuring they remain manageable by the time they reach the Flatten layer. Our input images will initially be 150x150 pixels, and as they progress through the network, the feature maps will gradually decrease in size to 7x7 before reaching the Flatten layer. This input size is somewhat arbitrary but fits our specific requirements well.

Training samples	**Validation Accuracy**	**Test Accuracy**	**Data Augmentation**
1000	78%	69.4%	NO
1500	83%	79.5%	YES
2000	86%	84.3%	YES

PRE-TRAINED MODELS





Conclusion:

The model configurations, along with the sizes of the training, testing, and validation datasets, are detailed in the preceding tables. We present results for the newly developed model both with and without data augmentation, as well as for models trained with larger training datasets or adjusted training and validation dataset sizes. This comparison includes metrics such as overall accuracy, validation accuracy, and the effect of data augmentation on the pre-trained model.

The analysis shows that models utilizing data augmentation do not consistently outperform those that do not. Additionally, increasing the size of the training dataset or altering the validation dataset size tends to enhance model accuracy. When comparing pre-trained models with data augmentation to those without, it becomes evident that data augmentation does not significantly boost either the model's accuracy or its validation accuracy. Overall, pre-trained models tend to perform better than those trained from scratch, especially in situations with limited training data.