Content Based Image Retrieval

Team members –

Bhavana Kurra

Vijayasimha Reddy Bheemireddy

Rachana Reddy Podduturi

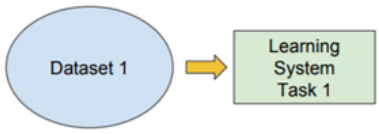
Problem Statement –

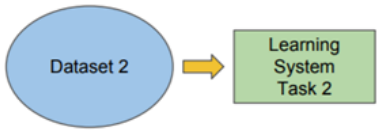
Building a Machine Learning model which displays the most relevant image in the user/local database for a given input image using transfer learning.

Background –

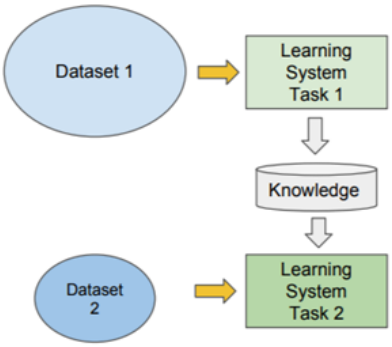
a) Traditional Machine Learning vs Transfer Learning -

Traditional Machine Learning creates a separate model for each dataset/task.

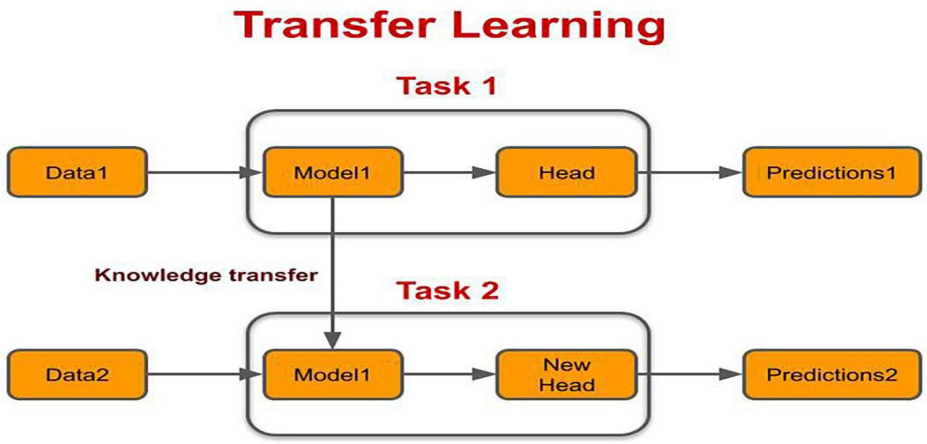




Transfer learning stores the knowledge gained while solving one problem and applying it to a different but related problem.



b) How are we using transfer learning in this project?

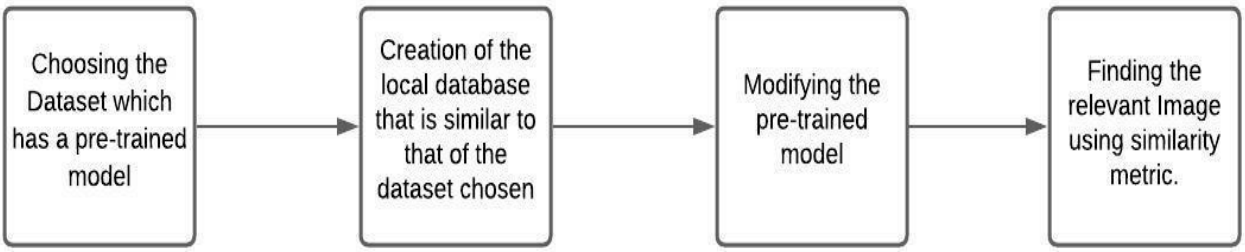


In the above mentioned image

1. The data1 is the ImageNet dataset.
2. The Model1 is VGG(Visual Geometry Group)-16
3. Task1 is ImageNet data classification
4. The data2 is the Database2(explained in the Methodology-step2) that we created.
5. Task2 is Image Retrieval from the Database1(explained in the Methodology-step2).

Architecture / Steps to be followed –

1. Choosing the Dataset which has a pre-trained model
2. Creation of the local database that is similar to that of the dataset chosen.
3. Modifying the pre-trained model
4. Finding the relevant Image using similarity metric.



Methodology –

Step 1 - Choosing the Dataset which has a pre-trained model

We have chosen the **ImageNet Dataset** and have chosen a **pre-trained model** over the ImageNet dataset.

a) Description about ImageNet Dataset –

- ImageNet is a dataset of over 15 million labeled high-resolution images belonging to roughly 22,000 categories.
- ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) uses a subset of ImageNet with roughly 1000 images in each of 1000 categories.
- Overall, there are roughly 1.2 million training images, 50,000 validation images, and 150,000 testing images

b) Classes in ImageNet Dataset -

ImageNet has multiple classes and subclasses. Some of them are -

- Goldfinch, cock, hen, ostrich, magpie, vulture, macaw etc are the classes that belong to the Bird Category.
- Tiger sharks, great white sharks, tench, goldfish, etc are the classes that belong to the Fish Category.
- Bullfrog, Tree frog, Tail frog, etc are the classes that belong to the Frog Category.
- Loggerhead, leatherback turtle, mud turtle, etc are the classes that belong to the Turtle Category.

c) Pre-Trained models over ImageNet –

1. AlexNet
2. VGGNet
3. Inception
4. ResNet
5. Xception etc

Step 2 - Creation of the local database that is similar to that of the dataset chosen.

a) Description about local database -

We have created a local database with a set of images from different categories that are similar to the ImageNet dataset but that don't exactly belong to the classes that exist in the ImageNet dataset. We have created a local database of 245 images. We have divided the local database into the two databases -

1. **Database1** - The database which has the images to which we extract and store the features using the pre-trained model for future usage.

2. **Database2** - The test database which has the images that we compare to that of the images in the Database1 to evaluate the pre-trained model performance.

The Database1 has 195 images in total and the Database2 has 50 images in total.

b) Image Categories that exist in the local database -

1. Aeroplane
2. Avocados
3. Bike
4. Bus
5. Car
6. Caravan
7. Carrot
8. Ceiling
9. Coffee
10. Guava
11. Hag
12. Helicopter
13. Hovercraft
14. Lungfish
15. Lychee
16. Motor
17. Mulberry
18. Onions
19. Pigeon
20. Plums
21. Rice
22. Sparrow
23. Spinach
24. Straightener
25. Trucks

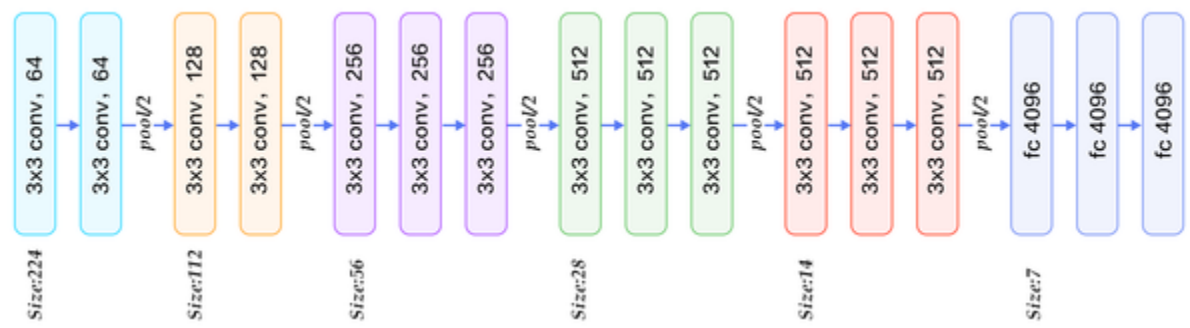
c) Need for the creation of a local database -

Our task is to retrieve the top 3 most similar images in our database(Database1) for a given query image(All the images in the Database2). That is, we created an Image Retrieval model using the knowledge gained by the VGG-16 model while trying to classify the ImageNet dataset. For the Image Retrieval model creation, we have created the Database1 and Database2.

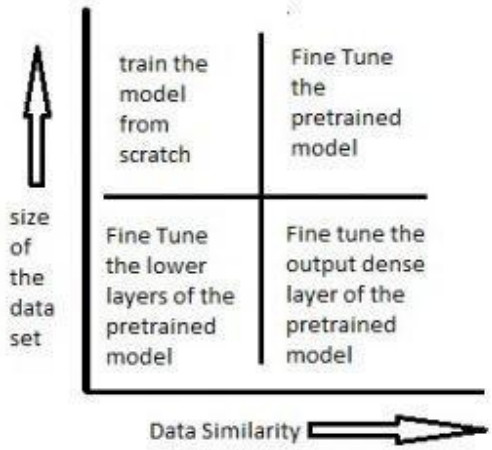
Step 3 – Modifying a Pre-trained model

We have chosen the VGG-16 model among the pre-trained models (AlexNet, VGGNet, Inception, ResNet, Xception, etc) that exist over the ImageNet dataset.

VGG 16 architecture –



In order to modify the pre-trained model, we are considering the local database similarity to that of the ImageNet dataset and the size of the ImageNet dataset. And based on the below image, we modified the pre-trained model.



Since our local database images are similar to that of the ImageNet database and also the ImageNet dataset has roughly 1.2 million training images, 50,000 validation images, and 150,000 testing images – we have opted for the 'Fine Tuning the pre-trained model either by the output dense layer or by the higher dense layers' option.We have used the higher dense layer(i.e the layer before applying the softmax function) to extract features.

Step 4 – Finding the relevant Image using similarity metric

We need a similarity metric to retrieve the relevant images from our local database(Database1) for a given query image. We have used the "**Cosine Similarity"** as the similarity metric inorder to compare the query image(i.e, Any image from the Database2) to that of the images for which we have extracted and stored the features(i.e, All the images in the Database1).

Result -

There are 50 images in the Database2 which we used as query images. Since, for a given query image we get top 3 relevant images we got 150 images as a result of querying 50 images. Out of 150 results, the model has retrieved 11 images that don't belong to the query image class i.e, 139 images are correctly retrieved. Therefore, the accuracy of this model is **92.66%** i.e., (139/150)*100.

Summary -

In Summary, we extracted and stored the features for all the images in the Database1 using the VGG-16 model(The extracted features are stored in the featureCNN1.h5 file). Later, we extract features 'f' for a query image(Any image from the Database2) and compare 'f' to all the images(all the features) in the Database1 inorder to find the top3 relevant images. Here we have used all the images in the Database2 as the query images and stored the results (results are stored in the Results.csv file) from which we calculate the accuracy of our model.