Duplicate Image Detection

T.Vineeth Kumar 15CS10047
P.Hari Krishna 15CS30022
D.Akhil 13CS10018

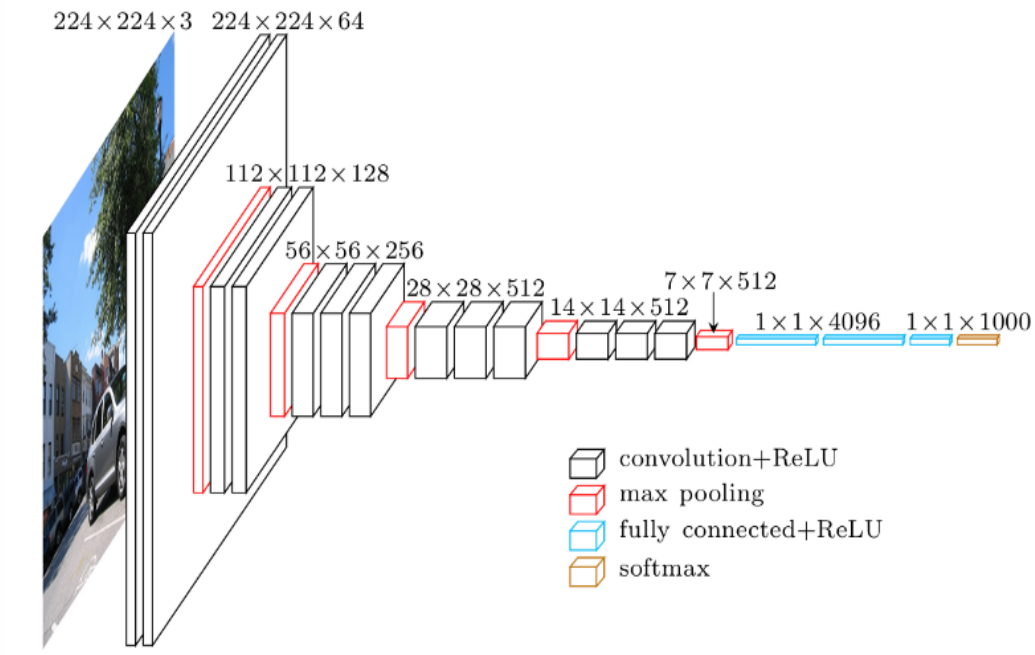
Group No. 37

1 Problem Statement

From a set of heritage images in duplicate image dataset, use any image as query and detect all the other duplicate instances of its presence using image matching by features (any features) and report the performance of retrieval.

2 Design of the Application

- **VGG** is a convolutional neural network model for image recognition proposed by the Visual Geometry Group in the University of Oxford, where **VGG16** refers to a VGG model with **16** weight layers.



- Using libraries, **VGG16** feature vectors for all the images and the query image are found.
- The image corresponding to the nearest VGG16 vector from the query VGG16 vector is reported as the duplicate image.

3 Libraries used

(a) keras

- keras.preprocessing
- keras.applications.vgg16

(b) numpy

4 Datasets used

- 10 sets of images, each set containing 4 similar images are selected. They are divided into two pools of 10 and 30 images.
- For each image from the pool containing 30 images, duplicate image is found from the pool containing 10 images.

5 Results

For all the images from the pool containing 30 images, correct duplicate images are found.