

# Advanced Machine Learning - Assignment 4

Pranav Kasela  
846965

## Introduction

The dataset consists of images ( $150 \times 150$ ) with 6 balanced classes: Forest, Mountains, Buildings, Sea, Glacier and Street (Figure 1). The data is divided in 14034 training samples and 3000 test samples, the training set is splitted into training and validation (80%-20%) during the training to test different hyperparameters.

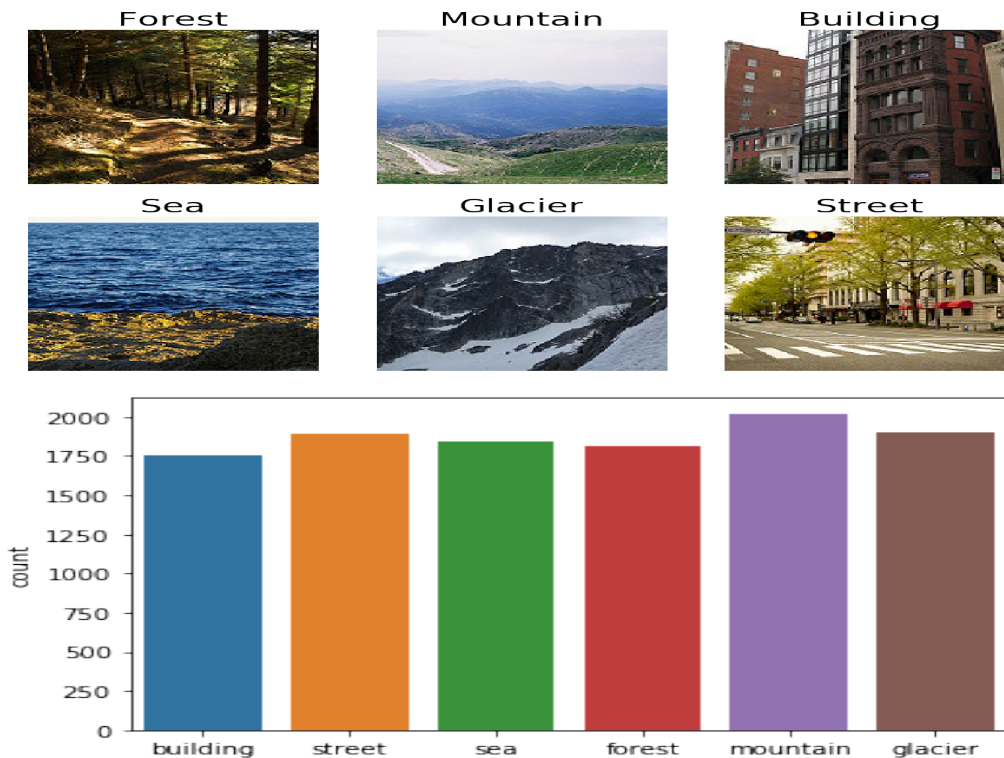


Figure 1: Example and distribution of the 6 classes.

## Models

For the assignment the model used for transfer learning is the VGG16, the part of it's architecture without the final FC layers is represented in Figure 2 along with the points where the four cuts are done, from the Figure 2 also the details about the layers can be deduced. The features are extracted with a GlobalMaxPooling layer in all four cuts, so number of the features are the number of filter in the last layer considered, this was done to avoid having a lot of features during the training another approach could have been using PCA or similar techniques before training the classifier. After the feature extraction the machine learning model used is the SVM with linear kernel in all four cases.

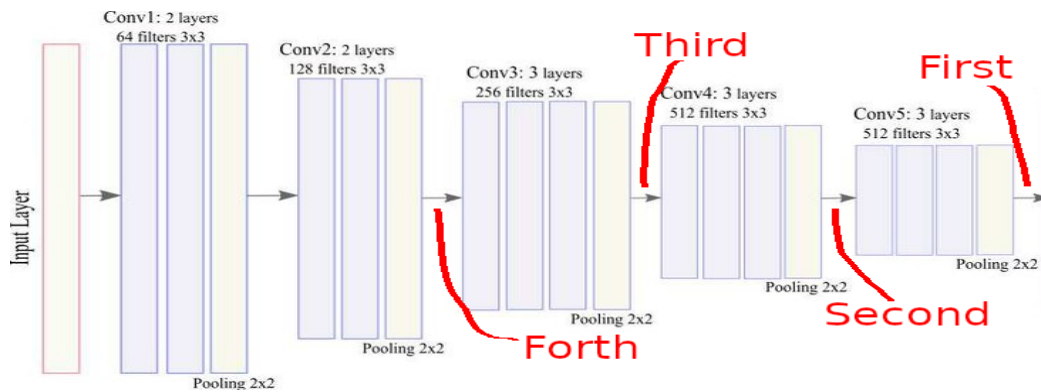
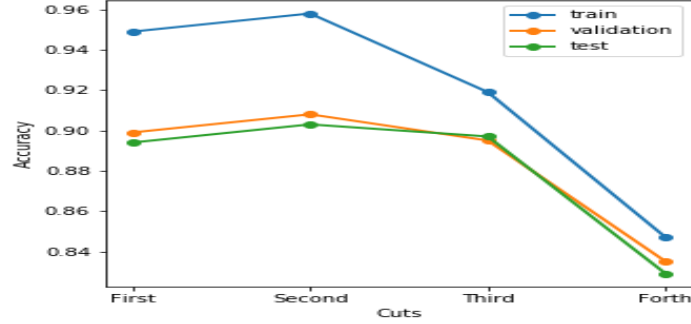


Figure 2: Top of VGG16 Architecture and the cuts for Transfer Learning

## Results

Both in the First and Second cut the feature extracted are 512, while in the Third and Fourth cut there are 256 and 128 features respectively, these cut are done at the end of each block (consisting of a few Convolution followed by a Pooling layer) of the VGG16 to see how much the classic machine learning classification models are able to distinguish the images based on the the features obtained by VGG16 pretrained on imagenet dataset.

Here are the accuracy results using the Support Vector Classifier with linear kernel, the C coefficient was optimized on the validation set in each cut, so it assumes different value in each cut:



	Train	Validation	Test
First Cut ( $C = 0.5$ )	0.949	0.899	0.894
Second Cut ( $C = 1$ )	0.958	0.908	0.903
Third Cut ( $C = 0.5$ )	0.919	0.895	0.897
Forth Cut ( $C = 0.1$ )	0.847	0.835	0.829

From the values reported in the table and from the accuracy plot the best cut seems to be the Second one with an accuracy of  $\approx 90\%$  in validation and test. In the Forth cut the performance got drastically worse, losing about 11% of accuracy in training and 7% in test and validation compared to the Second Cut. After observing such a deterioration in accuracy no further cuts were tried, assuming it will not improve.

The result could have been expected since there are some classes very similar to the imagenet dataset for example sea-(seashore, coast), street-(street sign), so going deeper in the model will create more interesting features for the classification.