# Intel Scene Classification Dataset

There are approximately 25,000 color images which depict natural and urban scenery in six different categories in the Intel Scene Classification Dataset. The six categories are: Buildings, Forest, Glacier, Mountain, Sea, and Street. The images are almost all uniformly sized at 150 by 150 pixels, with some discrepancies. The dataset was originally released by Intel for a challenge on Analytics Vidhya image classification competition. Ever since, the dataset has been used in some cases as a benchmark for evaluating scene classification models. The images are already separated via a train – test split: the training set total has approximately 14,000 images, the testing set has about 3,000 images, and the validation set has 7,000 images approximately.

# Exploratory Data Analysis

The dataset is relatively balanced along the six categories, with counts ranging from approximately 2,000 to 2,500:

A graph of a number of blue bars

AI-generated content may be incorrect.

Specifically, the highest representation was with the Mountain category, whereas Sea and Buildings had the least representation.

We also used a library to scan all of the images in order to see if any of the images were corrupted, the files, but none of them were. Also, we checked to see if any had low pixel initensity variance:

A diagram of a normal deviation distribution with Ryugyong Hotel in the background

AI-generated content may be incorrect.

But even the images with the least still had a lot of discriminative information between the classes, so no data cleaning needed in terms of this:  
A collage of different images

AI-generated content may be incorrect.

We had to do a small amount of data cleaning regarding the dimensionality of the data. Most of the dataset was 150 by 150 pixels, but there were some images that were slightly off in that dimension which we excluded:

A white background with black text

AI-generated content may be incorrect.

We also plotted the histograms of the different color channel values across each of the cannels in order to see if any of them were particularly noisy:  
A screenshot of a graph

AI-generated content may be incorrect.

None of these particular color channels are particularly noisy, but the blue color channels do have the most noise and the least overall intensity.

# Motivation

Our goal then is to train DenseNet-121 model to predict persistence landscapes derived from cubical filtrations from these scene images. What we want to do is determine if the DenseNet-121 architecture that we train is actually learning the topological features (holes in cubical filtrations) of the image or if it is only using irrelevant details in order to predict this. For this reason, we will compare the SHAP values at each pixel from the DenseNet against the actual holes of the image that our persistent homology is using, and we will compare to see if Deep Learning is actually learning the Topology or just approximating an arbitrary vectorization (Persistence landscapes) without truly learning the underlying structure.