Pex Machine Learning Technical Challenge

Part 1 - Create a labeled image dataset with two classes:

1. Indoor photographs (e.g. Bedrooms, Bathrooms, Classrooms, Offices)
2. Outdoor photographs (e.g. Landscapes, Skyscrapers, Mountains, Beaches)

Download a subset of examples from the YouTube-8M labeled video dataset:
https://research.google.com/youtube8m/explore.html

Extract relevant frames from the videos to build a balanced dataset of indoor and outdoor images. The dataset should contain a few thousand images in total. This task can be performed with tools like OpenCV or FFmpeg.

Create a train/test split of the data.

Part 2 - Train an image classifier capable of detecting if a scene is indoors or outdoors:

The model can be an artificial neural network or another approach of your choosing.

Evaluate the accuracy using the ratio of **all true results / the total number of examples tested**.

Part 3 - Create an evaluation tool for classifying single images with the trained model:

This CLI tool needs to allow for image files to be input one at a time.

The output needs to return a string containing the predicted class label.

Requirements

- Build a solution using one or multiple of the following languages:
    - Python
    - Java
    - C
    - C++
    - Go
    - Bash
- Write at least one unit test.
- Find a solution within three days. The whole challenge should take about 5 hours to complete.
- Upload your code and data from the final trained model to Github.

We will be looking for well written code and proper data preparation techniques. The resulting accuracy metric will not be considered as import as how you design and handle the dataset.