ING Bank Śląski – Computer Vision, Object Detection, Image Classification

Description

As a financial institution, one of the key aspects of our operations are mortgages. The crucial factor in this process is the evaluation of a loan collateral. Currently this process involves multiple manual inspections, sometimes involving field trips and surveys in numerous institutions.

The main goal of this task is to automate the most time consuming process – evaluation of a loan collateral. Usually, in case of mortgages, the collateral is some kind of real estate. As you may guess, the major factor in determining the value of property is its state and technical condition. During this task you will have to deal with the problem of automatic interior state detections, as well as the technical condition detection.

Dataset description

Your dataset consists of multiple public photos of both the interior and exterior of apartments and houses. Dataset is divided into separate folders, each folder contains between …. and … photos belonging to a single decision class.

Each photo is a valid jpg, some of them may have watermarks. Pictures have different quality, some of them may be even black and white. People faces should not occur in these samples, but should you find any – please give as feedback about it.

Task 1 (40% of total score):

Having the same set of pictures, your task is to find objects of interests among these. For example, we ask you to find all the chairs or bathtubs. The closed set of labels (one label per line) is given in task1_labels.txt file.

Task 2 (40% of total score):

Having the set of classes (one class per line, file: task2_classes.txt), your task is to predict what kind of room type (or generally speaking – picture class) it belongs to. You may approach this task in multiple ways: as a single-label object detection problem (find exactly one room type per picture) or as a decision problem (having the output of Task 1, predict the class by looking at objects inside the room).

Task 3 (20% of total score):

Having the set of pictures, your goal is to predict interior state, as well as the technical condition of a room/apartment.

Let's assume that interior state is given by four values: **1,2,3,4** and the technical condition is also described by four classes: **1,2,3,4.**

Evaluation

In order to speed up the process of evaluation, please provide a csv file with all the answers. The sample output file is given in eval_sample_output.csv file.

Your answer should be given in the same format. Each row denotes one picture file. The first column is the full filename with extension. The second one is the interior state. The third column is the solution for task 2: it's a single expression denoting a single class taken from task2_classes.txt file. The fourth column is a technical condition. The following columns contain the results of Task 1. Each label possible to detect in pictures has its own column. The value of "1" means the analysed objects is found in the picture, "0" means that this object is not found. No other symbols are allowed.

You are required to give full answer for every picture in test set.

Your score will be evaluated by the f1 score of your answers to reference answers with weighted average of label count per task weight. Example scoring for task 1:

Answer:

fileName	Bathtub	Bed	Chair	Couch	Window
Answer0.jpg	1	1	0	0	1
Answer1.jpg	1	0	1	0	1
Answer2.jpg	0	1	0	1	0
Answer3.jpg	0	0	1	0	1
Answer4.jpg	0	0	0	1	1
Answer5.jpg	1	0	1	0	1
Answer6.jpg	0	0	0	0	0
Answer7.jpg	0	0	0	1	0

Reference:

fileName	Bathtub	Bed	Chair	Couch	Window
Answer0.jpg	1	0	1	1	1
Answer1.jpg	1	0	1	0	1
Answer2.jpg	0	1	0	0	0
Answer3.jpg	0	0	1	0	1
Answer4.jpg	1	1	1	0	0
Answer5.jpg	0	1	0	0	1
Answer6.jpg	1	0	0	0	1
Answer7.jpg	0	1	1	0	0

$$F1\ Score = \frac{2*Precision*recall}{Precision+recall}$$

$$Score_{Task1} = \frac{2*0.(6)*0.5}{0.(6)+0.5}*0.2 + \frac{2*0.5*0.25}{0.5+0.25}*0.2 + \frac{2*0.(6)*0.4}{0.(6)+0.4}*0.2 + 2*0*0.2$$
$$+ \frac{2*0.8*0.8}{0.8+0.8}*0.2 = 42,61$$

Final remarks

- Assume, that all text files given to you are UTF-8 encoded, with Unix style line endings. Your answer should be given in same fashion

- Inside data folder you will find the validation set. All the pictures inside it are labelled, but marked as belonging to a "validation" class. You can, but do not have to use this pictures to improve accuracy of your solution.
- You are given the "output_all.csv" file with labels and predictions for the whole train set.
- Inside "data" folder you will have all the pictures with respective folders matched to a label from Task 2.
- There's a src folder with a skeleton Python code for generating sample output which validates correctly
- Last, but not least, you have an "ingscorer" folder contain python module used for scoring.

Submission instructions

Submit once per 15 minut.

Additional task

Welcome to a marvellous world of real estate. All of our consultants are currently busy, so you'll have to help yourselves.

Inside "input" folder you'll find 3 sample test sets. Each set consists of multiple files:

- Jpg files with pictures
- Json files with tags describing each picture
- Summary file in which every picture is labelled to a correct room class

Your task is to prepare a brochure/flyer/offer. Make it graphical and modern. Use NLP and text generation methods to utilise metadata and provide a human readable (and well composed!) description of each test set (assume that each test set is a separate flat and you have pictures of all rooms).

Generate your answer into a PDF file (one page per each test set).

No cheating please! During evaluation phase you will get a folder with same structure as "input" directory. Your solution has to generate the mentioned PDF file on the fly.