

Name: Elshan Naghizade

Title: Python library to transform image datasets into query-able metadata tables.

1) What am I going to do?

I am planning to build a python library which would extract metadata from image datasets to transform them into tables and bring them into a SQL-like query-ready format. Given an image of a popular format (for instance, png or jpeg) this python module would extract useful features to eliminate the need to manually develop data-loaders for preprocessing. Those features might be (not limited to) size, edges, corners, blobs, color-maps, etc.

2) How is it done today? Current Limitations?

Transformation of any image dataset into a processable state is done manually by means of data-loaders written in an ad-hoc manner (PyTorch Dataloader Class). Moreover, each dataset requires manual transformation into python objects (numpy, pandas.dataframe) to perform even basic cleansing or sorting.

3) What is your idea to do something better?

Automatically extracting the features universal for most image datasets and transforming them into a structured form would allow using conventional SQL queries on those metadata tables significantly cutting down the time spent on developing case-specific dataloaders and, most importantly, bringing SQL's flexibility and speed. A further step that might be taken later is to test the generated metadata tables in scalable NoSQL databases.

4) Who will benefit from your work? Why?

Computer Vision engineers spend a significant amount of time to cleanse and transform their image datasets into a state suitable for the model to be developed. Having simple SQL queries to snap through the datasets would simplify the automation of data-preprocessing requiring less manual interventions into the dataset itself, while interacting with the metadata itself.

5) What risks do you anticipate?

Metadata generation is a tedious process complicated by determining the selection of features to be extracted and the way they will be represented in a table (For instance, edge inflections might be a tough feature to store since it doesn't have crisp coordinates on the image arrays).

Moreover, it is probable that the image file parser would be parallelized since data-level concurrency is easily achieved by dividing the dataset into groups of images which might introduce certain HDD write collisions when writing the extracted metadata onto storage which is not concurrent by nature (at least, in Windows)

6) Out of pocket costs? Complete within 11 weeks?

I don't think there will be much out of pocket costs since most of the development can be carried out on the free version Google Cloud Colab with certain caveats. Memory limitations of the VM will restrict the development to HD-size images along with a single-core CPU. A basic working demo is believed to be feasible in 11 weeks in my opinion.

7) Midterm results?

I am planning to build the "scaffold" of application (simple feature extraction, like width, height, colormap, interception-based cropping queries) and to test it with MySQL.

8) Final Demonstration?

Sobel and Prewitt edges, Harris corner detectors and some other sophisticated features to be made available within SQL queries on metadata tables.