MACHINE LEARNING
CAPSTONE PROPOSAL

Davi Sales Barreira
August 27th, 2017

1. DOMAIN BACKGROUND

In the Civil Engineering field, the use of aggregates is very common practice in many situations. Understanding their properties is therefore important due to possible effects that these may cause. One example is in the production of concrete asphalt, in which aggregates superficial texture are known to affect the stiffness of the concrete asphalt (Singh, Zaman e Commuri, 2012) and to be related with it’s Flow Number (Pazos, 2015), which is a parameter used to measure the rutting potential of asphalt concrete mixtures (Wisconsin Highway Research Program, 2013).

Measuring aggregates superficial texture is no trivial matter. In Brazil, there are still no standardized procedure for evaluating such property. To solve this problem, researchers have come up with different techniques. One of such techniques was developed by Masad (2005) and uses an equipment called Aggregate Image Measurement System (AIMS). This equipment captures high quality images of the aggregates surface and utilizes wavelet analysis to measure a *texture index* (MASAD, 2005). Such index is a continuous variable that is used to classify aggregates as shown in the table below:

Texture Index	**Aggregate Classification**
0 - 165	Polished
165 – 275	Smooth
275 – 350	Low Roughness
350 – 460	Moderate Roughness
> 460	High Roughness

Table 1 – Aggregate texture classification (MASAD, 2005)

Although the use of the AIMS equipment may solve the problem of measuring superficial texture for aggregates, such method requires researchers to have access to such equipment and to specialized technicians trained to operate it. Therefore, simpler methods for obtaining such texture index are imperative.

2. PROBLEM STATEMENT

Convolutional Neural Networks (CNN) have been extensively used with regards to image analysis, although more commonly in classification tasks. The aim of this project is to build a CNN that estimates the texture index of aggregates using their images. The texture index predicted is the same obtained through the use of AIMS, therefore, the goal is replicate the results obtained when using this equipment.

3. DATASETS AND INPUT

The dataset was provided by the Department of Transport Engineering (DET) of the Federal University of Ceara (UFC). It consists of 1425 images of aggregates (Figure 1) with their respective texture indexes. The texture indexes were obtained using the AIMS equipment. The aggregates analyzed were gathered from three quarries located in the state of Ceara, Brazil. The CNN model will use the images as input and the output will be the texture index.

The original images are 480 by 640 pixels with grayscale. For the model, the number of pixels in each image will be decreased, and the effects of image quality will be evaluated. Due to the dataset not being very large, data augmentation might be implemented by flipping the images with an effort to increase the sample size. Some of the images are errors (Figure 2) and will be removed from the original sample.

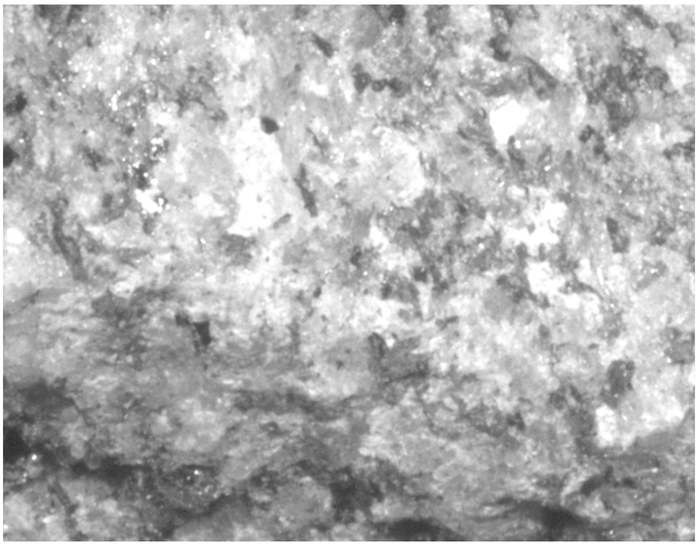

Figure 1 – Aggregate superficial texture

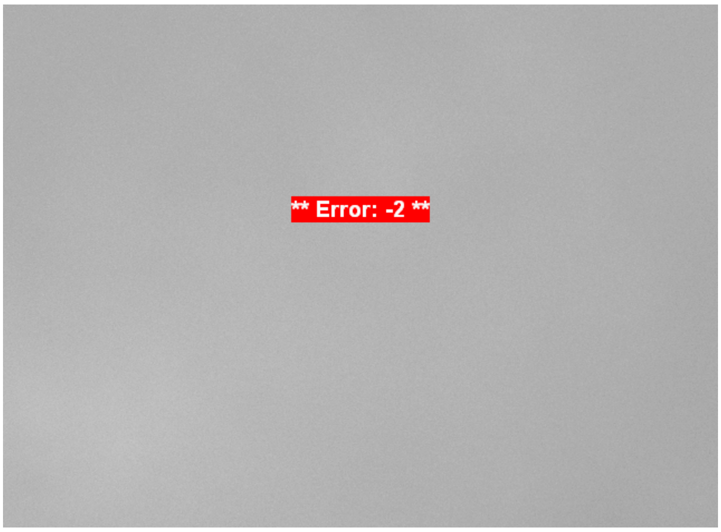

Figure 2 – Error image

4. SOLUTION STATEMENT

The project will implement CNN models that are usually used for classification task, and apply them in a regression problem. The objective is to obtain the texture index as close as to the one obtained through the AIMS equipment. The CNN model will be developed using Keras and Tensorflow as backend.

5. BENCHMARK MODEL

There are no approximate methods to obtain texture index, only using the AIMS equipment. Therefore, the benchmark model will be the average texture index for the test set.

6. EVALUATION METRICS

Two evaluation metrics will be used. The first one will be the R-Squared metric. This metric determines how well the model “explains” the variation in the data compared to using the mean (the benchmark). The goal is to obtain an R-Squared closer to 1 as possible.

The second metric will be the mean absolute error (MAE) of the model. While the R-Squared is a relative metric (performance compared to using the mean), the MAE is an absolute metric and it’s magnitude will be evaluated to see if the error in the model is small enough to give reliable results. It will be considered that the MAE must be smaller than 75 to be reliable, since this is the range size of the smallest classification bucket (Low Roughness, texture surface goes from 275 to 350).

7. PROJECT DESIGN

Programming Languages: Python 3.5

Libraries: Keras, Tensorflow, Sci-kit Learn, Matplotlib, Pandas, Pillow.

The workflow for the project will consist of first obtaining the dataset and organizing it so the images can be associated with the value of the texture indexes. The images will be imported using the Pillow library and have the quality reduced to 20% of the original size. The images that contain errors will be visually inspected and removed from the sample. After that, data augmentation will be implemented by flipping the images 180º and keeping the value of the texture index. The values in the pixels of the images will be normalized by dividing them by 255. Finally, the dataset will be divided into training set, validation set and test set. The training set will be 64% of the dataset, the validation will be 16% and the training set will be 20%.

An initial CNN model will be created in Keras with a structure based on the ones used for image classification (github.com/udacity/aind2-cnn/blob/master/cifar10-classification/cifar10_cnn.ipynb). Such model will be adjusted for regression instead of classification. This adjustment will consist of removing the final *softmax* layer. Figure 3 shows the model structure to be implemented in Keras.

```
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 96, 128, 48)       816
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 48, 64, 48)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 48, 64, 96)        18528
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 24, 32, 96)        0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 24, 32, 192)       73920
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 12, 16, 192)       0
_________________________________________________________________
dropout_1 (Dropout)          (None, 12, 16, 192)       0
_________________________________________________________________
flatten_1 (Flatten)          (None, 36864)             0
_________________________________________________________________
dense_1 (Dense)              (None, 500)               18432500
_________________________________________________________________
dropout_2 (Dropout)          (None, 500)               0
_________________________________________________________________
dense_2 (Dense)              (None, 1)                 501
=================================================================
```

Figure 3 – CNN Model Structure in Keras

The model will be trained using several epochs and evaluated against the validation set using the R-Squared as the metric. The model that obtains the best result in the validation set will be compared to the test set. The R-Squared and MAE will be calculated in the test set to see if the model produced good results.

Based on the results of this initial model, different networks architectures will be tested by adding or reducing layers, changing kernel sizes and adding or removing dropout and pooling layers. This models will be compared in terms of R-Squared with their respective validation sets to see which perform better. After this, a final model will be chosen and again evaluated with the test set.

Finally, the effect of image quality will be studied. The original images will have the amount of pixels reduced to 50% and to 10%, instead of the original 20%. Both these datasets will then be used to train the previously chosen CNN model. Finally, the R-Squared and MAE calculated in the test set will be compared for the three datasets (10%, 20% and 50% reduction) to see the impact of image quality in the estimation of surface texture.

REFERENCES

MASAD, E. A. Aggregate Imaging Measurement System (AIMS): Basics and Applications. Texas Transportation Institute. The Texas A&M University System. Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration. Report no FHWA/TX-05/5-1707-01-1, 2005.

PAZOS, A. G. Efeitos de Propriedades Morfológicas de Agregados no Comportamento Mecânico de Misturas Asfálticas. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015.

SINGH, D.; ZAMAN, M.; COMMURI, S. Inclusion of Aggregate Angularity, Texture, and Form in Estimating Dynamic Modulus of Asphalt Mixes. Road Materials and Pavement Design, v. 13, n. 2, p. 327-344, 2012.

WISCONSIN HIGHWAY RESEARCH PROGRAM. Flow Number as a Discriminating HMA Mixture Property - No. 0092-09-0. Wisconsin, 2013.