List of group members:

- Tony Joo
- Ahmed Ibrahim
- Jordan Wu
- Hossein Kabir

Informative project title:

Synthetic data generation for Mask R-CNN enabled water absorption estimation in cement pastes

Project description and goals:

Durability of construction materials is mainly controlled by the rate of absorption of water by cement pastes. Specifically, most deterioration mechanisms in concrete structures occurs as a result of water intrusion into the cement paste matrix. Therefore, it is necessary to estimate absorption of water as a function of time, in which only one surface of the specimen is exposed to water. To do the experiment, an exposed surface of the specimen has to be immersed in water where the water ingress of unsaturated cement paste is dominated by capillary suction during initial contact with water. Originally this test was being operated by technicians in accordance with ASTM C1585 standard test method [1] such that the sample has to be weighed at specific time intervals to estimate absorption with time. However, the traditional approach is not only labor intensive and expensive, but also fails to provide the user with accurate results since the rate of absorption could not be measured in real time.

As a result, this course project aims to leverage computer vision to continuously estimate the water absorption level in cement pastes as a tool to indirectly measure the rate of absorption. For this reason, our group is going to analyze images taken from microscopy camera that was used to be placed in front of the sample (subjected to absorption test) and automatically extract water level with time. We realized that Mask Regional-Convolutional Neural Network (Mask R-CNN) method [2] is an excellent choice to be leveraged for analyzing image data to identify water level from the cement paste specimens. This method couples a deep learning Faster R-CNN with fully convolutional networks (FCN) to delineate regions of interest (ROIs) and classify every pixel of images. In this pixel-level water-level estimation algorithm, the user is required to <u>manually</u> annotate the dataset to make the objects recognizable for machine training. However, manual annotation of a large dataset is a time-consuming process, and more importantly, the boundaries of the image data may not be perfectly aligned, especially on complex geometries [3].

To showcase the subjectivity of image annotation, our group members (i.e., Hossein, Tony, Jordan, and Ahmed) did manually annotate few images by CVAT (i.e., an open-source annotation tool [4]) and realized that there is a notable difference between manual annotation of same image if done individually. Specifically, **Fig. 1** shows 3 images (cement pastes at different water levels) that are annotated by our group members and the associated error map (i.e., the standard deviation of the pixels between the segmented masks of each row) is shown in the last column denoting the remarkable subjectivity of manual image annotation. Consequently, there is a need for <u>automatic</u> image annotation specifically for larger datasets in order to accurately train the Mask R-CNN model. For this purpose, we will synthetically generate 2D models so we can automatically mark and segment the ROIs for accurate training of the Mask RCNN algorithm (time permitting, we will further generalize our model to create 3D models as well). To generalize our model, the synthetic image data will be made at different lighting conditions, color textures, water level orientations (e.g., horizontal, and inclined), and light reflections, see **Fig. 2**. At the end of this project, we will determine and document whether synthetic models can partially (i.e., combination of real and synthetic image dataset) or fully replace natural imagery for accurate measurements of water absorptions in cement pastes.

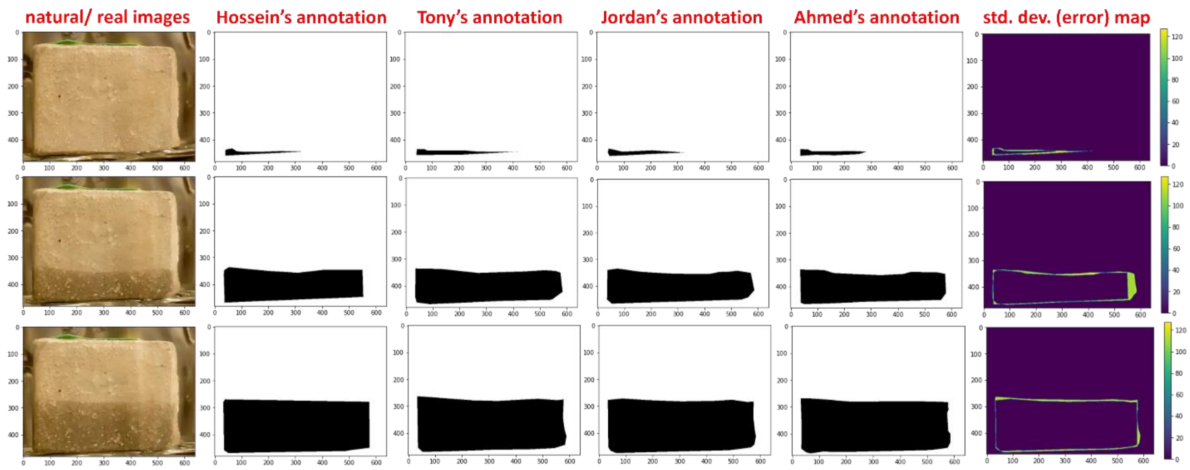

Fig. 1: Subjectivity of manual image annotation, which is showcased by our group.

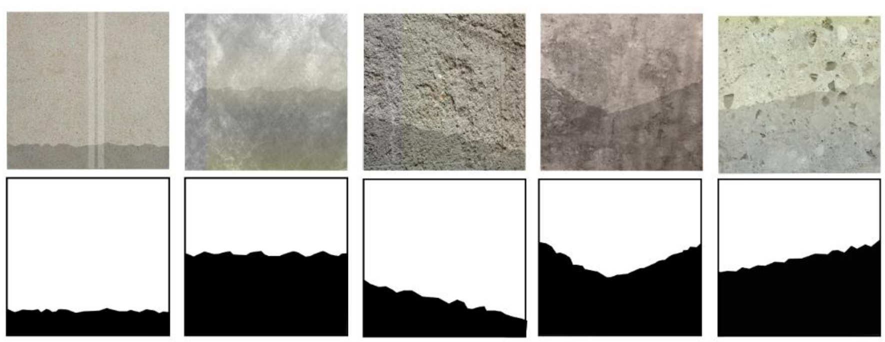

Fig. 2: Synthetic data generation to accurately measure water absorption

Member roles:

- Tony, Jordan, & Ahmed Ibrahim will do:
  - Manual annotation of natural/ real images.
  - Optimize the Mask R-CNN algorithm to be suited for our project.
- Hossein:
  - Generate synthetic images at different water levels, color textures, lighting conditions, and reflections.
  - Optimize the Mask R-CNN algorithm to be suited for our project.

Resources:

To write and implement the code, we will be extensively using the original Mask R-CNN source code built on FPN and ResNet101, which was released in 2017 and is currently under MIT License [5]. Currently we have approximately 1000 natural images that were previously taken by the microscopy camera (some of them are shown in the most left column of **Fig. 1**) which are going to be manually annotated by CVAT. As well, we have 250 synthetic 2D models, and we are going to make another 250 synthetic 2D models to generate approximately 2000 images (i.e., 4 different illumination conditions for each image) for training our model, see **Fig. 2**. We are also going to upload our images and their associated segmented masks to google drive and we will run the code in google Colab (our code is not computationally intensive and as a result the free version of Colab service should be sufficient for our project).

Reservations:

We expect to annotate 1000 real images for training the mask RCNN model (each of us will separately annotate the same 1000 images). The manual image annotation is expected to be a time consuming and tedious task but would hopefully be a good backup to evaluate the accuracy of our model. Also, we have never used synthetic data to train a deep learning model, and this experience would be the most difficult part of our project that we aim to work hard to accomplish. Furthermore, making the Mask RCNN model compatible with our own dataset is another challenge that we need to address by selecting proper model constants.

Relationship to your Background:

Jordan is a 1st year Master of Engineering student in Computer Engineering who has taken some undergraduate machine learning courses and have an understanding of the basic concepts and models. He is also familiar with pytorch to build recommendation systems, but he is new to computer vision.

Tony is a 1st year Master of Engineering student in Computer Engineering, and he has taken CS 498 DL in the previous semester and has worked on a project to develop CNN models for object detection. The image segmentation and working with synthetic images are new to him.

Hossein is a 2nd year PhD student in Civil Engineering, with some background in deep learning and computer vision. His PhD thesis is partly devoted to minimizing subjectivity in laboratory measurements. He has previously worked with Mask RCNN and realized how difficult it can be to annotate ROIs with complex boundaries. That is why he plans to leverage synthetic data to train Mask RCNN model for the first time as a project for CS 543 course.

Ahmed is a 1st year PhD student in Civil Engineering with some background in deep learning. His research is focused on leveraging machine learning for crack and anomaly detection on civil structures. By doing this course project he hopes to enhance his skills needed for his PHD thesis.

References

[1] ASTM C1585-20 (2020). American Society for Testing and Materials, Standard Test Method for Measurement of Rate of Absorption of Water by Hydraulic-Cement Concretes, Pennsylvania: ASTM International.

[2] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

[3] Li, S., & Zhao, X. (2021). Pixel-level detection and measurement of concrete crack using faster region-based convolutional neural network and morphological feature extraction. Measurement Science and Technology, 32(6), 065010.

[4] https://github.com/openvinotoolkit/cvat

[5] https://github.com/matterport/Mask_RCNN