American Society of Civil Engineers (ASCE) Computing Division
Visualization, Information Modeling, and Simulation (VIMS) Committee
and
International Association for Automation and Robotics in Construction (IAARC)
Workshops and Summer Schools Committee

3rd Annual VIMS / 1st VIMS-IAARC Joint Datathon Competition 2022

The VIMS Committee invites all interested parties to its 3rd Annual VIMS Datathon competition in conjunction with the 2022 International Symposium on Automation and Robotics in Construction (ISARC) conference that will be held in Bogota, Colombia. This year the competition is jointly organized and held with the IAARC Workshops and Summer Schools Committee.

Introduction: With the increasing awareness of sustainable development, reusing and recycling waste is gaining importance. Research has shown that a lot of material in construction and demolition waste can be effectively used in new construction. Can such material be automatically identified through computer vision and segregated using robots? As a first step, in this Datathon we attempt to find out what is possible through computer vision. Therefore, the goal is to identify recyclable objects such as PVC pipes and rebars within construction debris through image processing and artificial intelligence. The participants are required to select appropriate machine learning algorithms and tools for the task. The training data will be released to the registered teams prior to the start of the Datathon. Using this data, the team should develop a machine learning model that can be used to predict whether and where specified objects are present in images. The test data will be released to the teams on the final day of the competition.

Team Formation: Teams will consist of 4 or 5 undergraduate and/or graduate students, who have a vested interest in ASCE Computing Division and IAARC, machine learning, computer vision, and 3D data. Teams can be formed with students from the same institution or from multiple institutions. Each team must have a mentor, who is selected by the members of that team from full-time faculty members in one of the participating institutions. The faculty mentor should be a current member of any of the four ASCE Computing Division Committees (VIMS, DSA, EDU, Global) or IAARC. Only teams that are formed and registered through Google Form: https://forms.gle/PZ5JTjjc4C9M9zqf7 on or before **11:59 p.m. EST on June 20, 2022,** will be eligible to participate in the 2022 Datathon.

Benefits and Prizes: Winners will be announced and formally recognized **(with monetary prizes and other benefits)** during an online meeting held at the end of the Datathon period on **July 5, 2022**. Other attractive prizes will be determined based on the funding from the industry sponsors.

<u>*First place:*</u> $500 + free registration for the 2023 ISARC conference for one person from this team
<u>*Second place:*</u> $300
<u>*Third place:*</u> $200
In addition, all the participating teams will have the right to utilize the generated work and initial data in the future. The winning teams will be also acknowledged at the 2022 ASCE VIMS Committee's annual meeting.

Dataset: A set of around 500-600 images captured from new construction and demolition sites will be provided to all registered teams. Each image has a set of construction objects/materials that can be recycled and a corresponding description that provides the label of the materials seen on that image. The dataset captures the following **types of objects in general: rebar, bricks (full or broken), PVC pipes, plastic wires,** and **cementitious debris**. Access to the dataset will only be given to the registered teams. Teams may decide to augment or supplement the initial dataset for improving training accuracy. Sample images are provided at the end of this announcement. Teams will also utilize state-of-the-art labeling processes to label the initial dataset. Teams are expected to detect the silhouettes of objects in images.

Datathon Tasks: Teams will utilize the released dataset to train machine learning models for automated segmentation of objects for recycling processes. Teams are not constrained to use any specific software packages or platforms; however, they should be able to share their workspace during the online meeting for the final evaluation of their trained models by the judges. Teams can choose to use any suitable pre-trained models to start their training process. **<u>Specific tasks</u>** required from the teams are listed below:

- Use the released dataset (composed of ~500-600 images to be provided to registered teams via a Google Drive link once they register) to label **(1) rebar, (2) bricks, (3) PVC pipes, (4) granite, and (5) cementitious debris** as materials that have a potential to be recycled. The method of labeling is up to the teams. Teams may further do image augmentation steps to increase the data size. The focus of this Datathon is only on the 5 object labels given in this Datathon description. *Note: An image might contain multiple types of labeled objects and images are of different sizes and resolutions.*
- Use this labeled (and augmented) dataset to train machine learning models (algorithms of your choice or pre-trained models of your choice as the baseline models from github libraries) for the categories of objects given.
- While training your models, teams are expected to use an 80-20% split to train and validate their models. Teams are expected to release the results of their training effort by the interim report submission date (see below for important dates).
- The judges will release a test dataset (separate from the initial dataset) on the final evaluation date (July 5, 2022) and ask teams to download the testing dataset and use it during the online evaluation to show their testing results using IoU (intersection over union) as the metric. Teams will show IoU per label and mean IoU (mIoU) across all four labels.

Deliverables: Teams will provide their results in two parts.
  (1) <u>*Interim evaluation deliverable:*</u> The first part used for interim evaluation is due by **11:59 p.m. EST on June 28, 2022**, where the teams **will provide a set of PowerPoint slides** on (1)

the methodology they followed for the labeling process, (2) the algorithms they evaluated and finally chose for training their models, (3) the accuracy of their training results using IoU per label and mean IoU (mIoU) across all the four labels, and (4) a creative list of proposed ideas for how to leverage the dataset and the training models in research and practical settings. This interim submission should be uploaded to the designated Google Drive folder (link will be provided to registered teams). The submission deadline is on or before **11:59 p.m. EST on June 28, 2022.**

(2) Final evaluation deliverable: No deliverable will need to be submitted. The final evaluation will be online using the testing data released on **July 5, 2022**. Participating teams must attend the online meeting to be held at **10:30 am EST on July 5, 2022** to show their testing performance on the spot. The performance of the machine learning models will be evaluated using the IoU metric calculated per label and averaged across the labels.

Important Dates:

- Call for participation: **April 4, 2022**
- Team formation and registration deadline: **11:59 pm EST, June 20, 2022** (Registered teams will receive additional information on the next steps)
- Datathon period: **June 24 - July 5, 2022**
- Interim report submission: **11:59 pm EST, June 28, 2022**
- Decision on teams to continue to the final stage: **June 30, 2022**
- Release of the test dataset to teams: **10:30 am EST, July 5, 2022** (testing data can be downloaded starting at 10 am EST on that date but will be password protected until that team's turn for evaluation during the online meeting). Teams will receive testing data on the spot sequentially and will have 5 minutes to run their models and show the testing results.
- Final evaluation by the judges: **10:30 am EST, July 5, 2022**
- Final announcement of results: **July 5, 2022, by the end of the meeting**

Evaluation Criteria: VIMS Officers and IAARC Workshops and Summer Schools Committee (and ad-hoc experts, as needed) will evaluate all submissions, and identify the top 3 teams for this Datathon. Winners will be announced at the end of the online meeting on July 5, 2022. The evaluation criteria are listed below:

Interim Evaluation (30%) *[delivered by teams on June 28, 2022]*
1. Labeling process (10%): Methodology explained in the interim submission for the labeling process.
2. Training/Validation effort and accuracy (10%): Number and suitability of models trained, tuned, and evaluated before the final model is selected and submitted in the report. Performance of the trained models. IoU and mean IoU will be the main accuracy metrics.
3. Creativity (10%): Proposed list of ideas by the teams and their novelty (not thought of or identified in literature yet).

Final Evaluation (70%) *[delivered by teams during the online meeting on July 5, 2022]*
4. Model testing performance on the spot (70%): Teams will download the testing data to be released on the evaluation date (July 5, 2022) and have access to them when their team is

up for evaluation on that day. Each team will receive a unique password-protected link to download the testing data right before the online evaluation time. Passwords to open the downloaded dataset will be released to the teams when a team is due for the testing performance. This will be a process to make sure each team has the same amount of time reserved for testing. IoU and mean IoU will be the main evaluation metrics for this part.

Note: This Datathon will be held in a virtual mode. The participating teams do not have to attend the 2022 International Symposium on Automation and Robotics in Construction (ISARC) conference in person.

Disclaimer: ASCE VIMS Officers and IAARC Workshops and Summer Schools Committee reserve the right to interpret the rules of the competition.

ASCE VIMS Officers: Semiha Ergan, NYU (VIMS Chair); Fei Dai, WVU (VIMS Vice-Chair); Jing Du, UFL (VIMS Secretary); Reza Akhavian, SDSU (VIMS Member-at-large); Amir Behzadan, TAMU (VIMS Past Chair).

IAARC Representative: Benny Raphael (I.I.T. Madras)

Sample image1: Broken bricks

Sample image 2: Cementitious debris

Sample Image 3: Plastic wire

Sample Image 4: PVC pipes

Sample image 5: Rebars