AUTOMATIC GENERATION OF PHOTOREALISTIC TRAINING DATA FOR DETECTION OF INDUSTRIAL COMPONENTS

Yu-Hui Lee[1], Chu-Chun Chuang[1], Shang-Hong Lai[1], Zih-Jian Jhang[2]

[1]National Tsing Hua University, Taiwan, [2]Industrial Technology Research Institute, Taiwan

INTRODUCTION

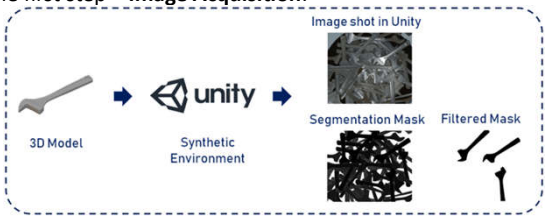
We propose a method to **automatically generate realistic training data for industrial components detection**.

- Our method can generate a large scale of various **synthetic images** associated with the corresponding **instance segmentation masks** through the concept of *domain randomization* and *style transfer*.
- Our method can enhance the performance of the wrench detection task obviously, and it can be easily extended to the detection of different kinds of industrial components.
- The first work that can synthesize realistic images and the associated segmentation masks for massive amount of piled industrial components.

SYSTEM OVERVIEW

We propose a **two-step** data synthesis method to generate a large scale of realistic images of piled industrial components.

- The first step – **Image Acquisition**:



- The second step – **Photorealism**:

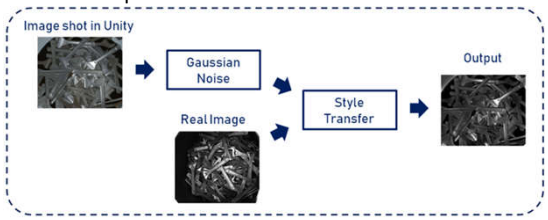


IMAGE GENERATION

1. Construct a **virtual environment** according to the real scene, included a *bucket*, *point light*, *60-80 wrench models*, and a *camera* facing the bucket.
2. Cast wrench models with different position and rotation randomly into the bucket every time an image is synthesized.
3. Randomly modify the *color* and the *metallic strength* of each wrench model in the scene.

Examples of real wrench images

INSTANCE SEGMENTATION MASK GENERATION

- For each synthetic image of wrenches, we generate the associated **segmentation mask including all of the wrenches** in the image.
- We also produce a **filtered mask** which preserves **only wrenches on the top layer** of each pile of wrenches.

Synthetic Images **Full mask** **Filtered mask**

- We cast a small number of *rays* from different positions of the wrench to the sky to check if the wrench is hit by other wrenches.

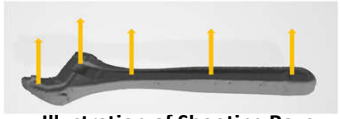
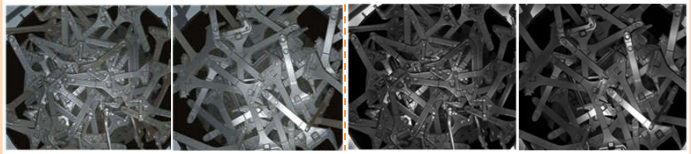


Illustration of Shooting Rays

PHOTOREALISM

- Applying **Gaussian noise** and **style transfer** to the synthesized images to make them more realistic.
- We adopt *Fast Photo Style[1]* for the style transfer task, and we randomly pick one real images of wrenches as the style image.

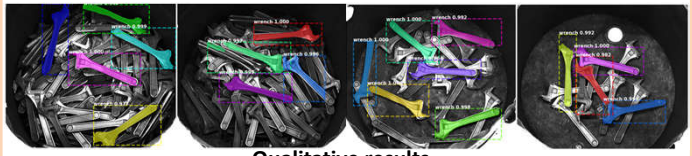


Left: before photorealism, Right: after photorealism.

[1] Y. Li, M.-Y. Liu, X. Li, M.-H. Yang, and J. Kautz, "A closed-form solution to photorealistic image stylization," arXiv:1802.06474, 2018

EXPERIMENTAL RESULTS

- We use *MASK R-CNN* as our wrench detector.



Qualitative results

Training Data	mask AP	bbox AP
Only R	57.1	72.0
Only S (w/o style transfer)	54.4	60.4
Only S (w/ style transfer)	59.0	64.7
R and S (w/o style transfer)	73.0	76.9
R and S (w/ style transfer)	**78.2**	**82.0**

Average Precision @0.5 IoU of using different training data for the wrench detection task. Real data(R): 392 images, Synthetic data(S): 1,000 images.

Training Data	mask AP	bbox AP
500 Synthetic images	71.4	74.6
1,000 Synthetic images	78.2	82.0
2,000 Synthetic images	79.8	80.2
4,000 Synthetic images	**82.0**	**82.9**

Average Precision @0.5 IoU of using both 392 real images and different numbers of synthetic images to train a wrench detector.