A Realistic Synthetic Mushroom Scenes Dataset

Dafni Anagnostopoulou, George Retsinas, Niki Efthymiou, Panagiotis Filntisis, Petros Maragos

dafnianagnostopoulou@mail.ntua.gr, gretsinas@central.ntua.gr, nefthymiou@central.ntua.gr, filby@central.ntua.gr, maragos@cs.ntua.gr

National Technical University Athens

The Dataset

The **Realistic Synthetic Mushroom Scenes Dataset** contains:

- **15,000 images** of synthetic realistic mushroom scenes.
- **2D annotations**: bounding boxes, segmentation masks etc.
- **3D annotations**: bounding boxes, 2D projections, 3D poses.

The dataset is generated through a **three-step pipeline** (https://github.com/dafniana/Synthetic-Mushroom-Dataset) and can be used to address **mushroom detection, instance segmentation, and 3D pose estimation** problems.

Figure 1. An image coming from the Realistic Synthetic Mushroom Scenes Dataset

Creating Synthetic Scenes

- **Ground:** Begin with a texture-less **square mesh**, add a small value of random noise, fill with triangle faces and assign a random **soil image as texture**. Using a RBFInterpolator **randomly deform** the ground mesh.
- **Mushroom Instance:** Begin with a **3D mesh** of an agaricus mushroom. Uniformly **scale** with a random scaling factor and smaller factors along each axis and **apply deformations** according to normals. Randomly **rotate** and **position** on the ground plane.
- **Mushroom Population:** Randomly **choose the number** of mushrooms in the scene. For each instance created **check for collisions with others**. Generate instances until the desired mushroom number is reached while discarding colliding ones.

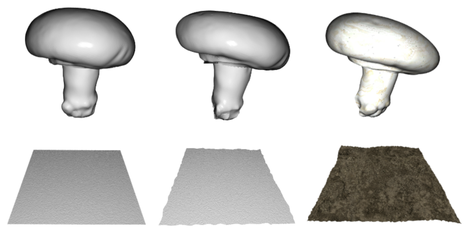

Figure 2. 3D mushroom and ground models used to build the synthetic scenes: initial models (left), deformed models (middle) and textured deformed models (right).

Rendering 3D Scenes to 2D Images

Randomly select viewpoints that look at the scene from **different angles** and **different distances** around it. Use the **Open3D visualizer [2]** to visualize the scene and capture the screen image. For every rendered image, render masks for all mushrooms visible from that viewpoint (Figure 3).

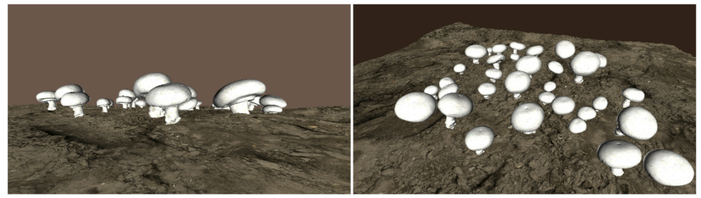

Figure 3. Rendered images of a synthetic scene from two different viewpoints.

From Synthetic to Realistic Images

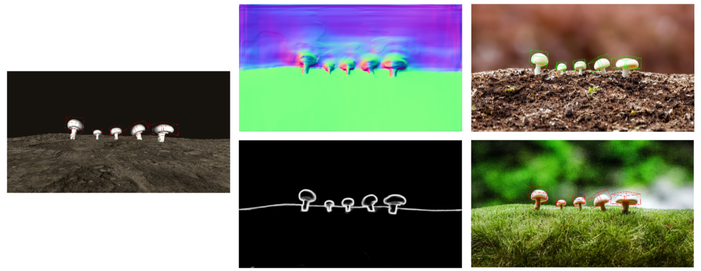

Figure 4. Retaining mushroom shape and pose using as control input to the diffusion model the normals maps (on top) and the HED boundaries (on the bottom).

Employ **ControlNet network [1]** to obtain realistic 3D annotated images.

Best control options to use as image based conditions are:

- normal maps
- HED boundaries

because they could **preserve the mushroom position, shape and pose information** (Figure 4)).

Experiment with directly using the rendered normal maps, which are more precise. Final result does not show any significant difference (Figure 5).

Create five realistic images per synthetic image using different prompts and control parameters.

Prompts						
white agaricus brown white dirty agaricus white agaricus with dirt white portabello portabello	mushrooms	on in	dark brown soil ground dark soil ground black soil ground brown soil dark soil black soil	and with	sky forest green forest trees garden farm	in the back background

Figure 5. Qualitative comparison between using directly rendered (top row) and estimated (bottom row) normal maps as control input to the diffusion model.

Exploiting the dataset

For **instance segmentation**, a Mask R-CNN model was fine tuned on our realistic synthetic data as well as only with the initial synthetic data.

Model achieves significantly better instance segmentation ability when trained with the realistic training data (Table 1).

Training Data	mAP	mAR
Synthetic	81.7%	86.4%
Realistic Synthetic	**89.5%**	**93.3%**

Table 1. Image segmentation results with simple synthetic training data and with the final realistic data.

The model was tested on images of white cultivated mushrooms, Agaricus bisporus, on industrial mushroom farms (Figure 6).

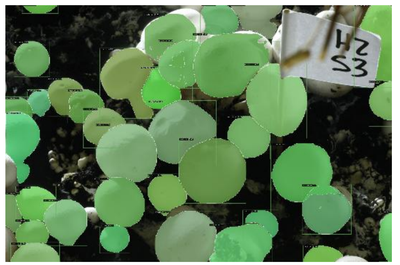

Figure 6. Example of instance segmentation on image taken from real mushroom cultivation bed.

Acknowledgements

This research has received funding from the European Union's Horizon 2020 research and 628 innovation programme under grant agreement no. 101017054 (project: SoftGrip).

References

[1] L. Zhang and M. Agrawala.
Adding conditional control to text-to-image diffusion models, 2023.

[2] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun.
Open3D: A modern library for 3D data processing, 2018.