Generating Diverse Agricultural Data for Vision-Based Farming Applications

Mikolaj Cieslak[1], Umabharathi Govindarajan[2], Alejandro Garcia[1], Anuradha Chandrashekar[2], Torsten Hardich[1], Aleksander Mendoza-Drosik[1], Dominik L. Michels[13], Soren Pirk[14], Chia-Chun Fu[2], Wojciech Palubicki[15]
[1]GreenMatterAI [2]Blue River Technology [3]KAUST / TU Darmstad [4]Christian-Albrecht University [5]Adam Mickiewicz University
www.greenmatter.ai www.bluerivertechnology.com

CVPR SEATTLE, WA JUNE 17-21, 2024

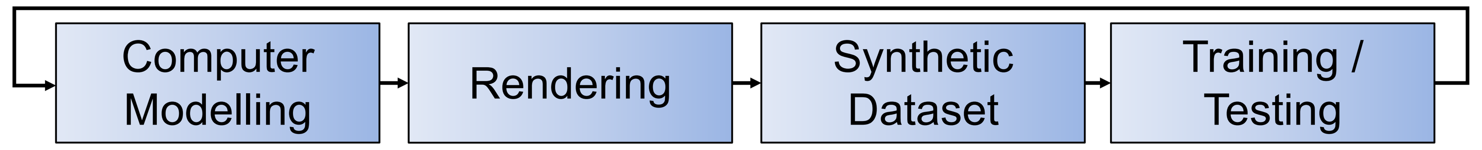
Introduction

Challenges of acquiring agricultural data for vision-based applications:
- varying crop stages, similar plant species, soil conditions, field arrangements

Advantages of synthetic data:
- cost-effectiveness and accuracy, diversity and edge cases, tailored conditions

Existing procedural methods are too broad in scope:
- lacking diversity in specific conditions (potential for overfitting)

Our contribution:
- a specialized, procedural approach for generating agricultural scenes

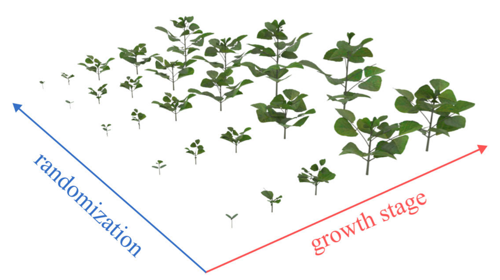


Procedural data generation

1) Plant modeling: we model different species and growth stages using an L-system-based modeling language in the Virtual Laboratory (www.algorithmicbotany.org) and with Geometry Nodes in Blender (www.blender.org).

2) Texture modelling: we obtain diffuse and alpha maps from images, but we generate height, normal, and roughness maps procedurally.

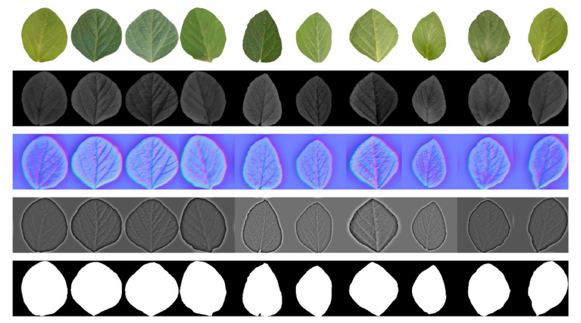
Soybean



Giant foxtail

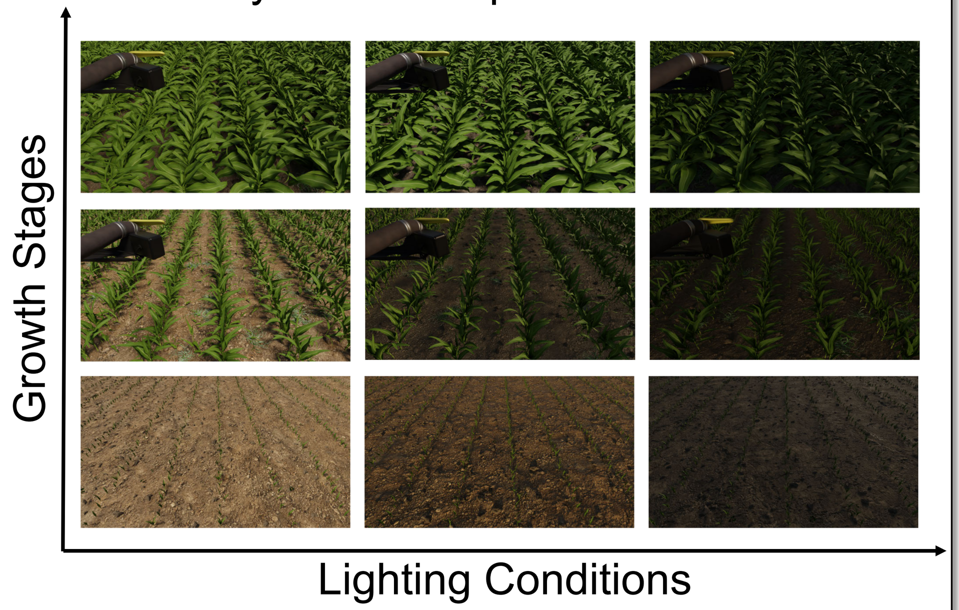
number of tillers

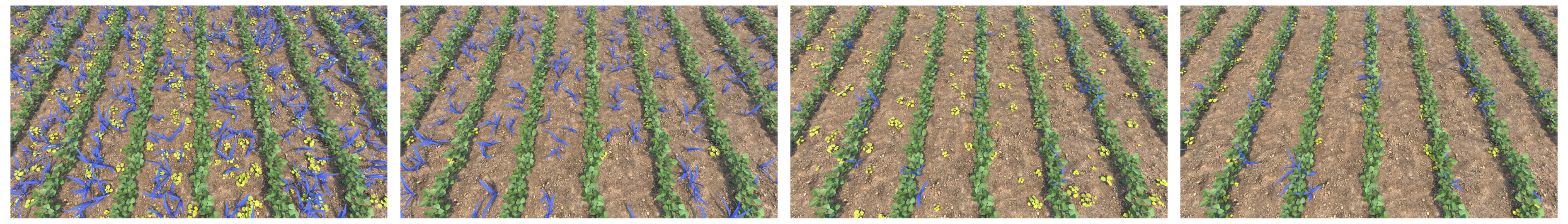
Soybean leaf textures



3) Soil modeling and field arrangement: we model soil color and moisture, and various distributions of plants in the field.

4) Image rendering: we use path tracing to generate photorealistic images, varying time of day and atmospheric conditions.





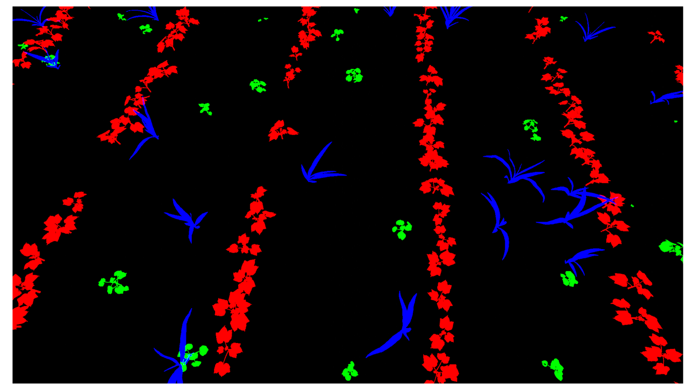
inter-row weeds ⟵⟶ intra-row weeds

Synthetic dataset

We generated a synthetic dataset of 12,000 images of soybean fields with labels for crops and weeds.

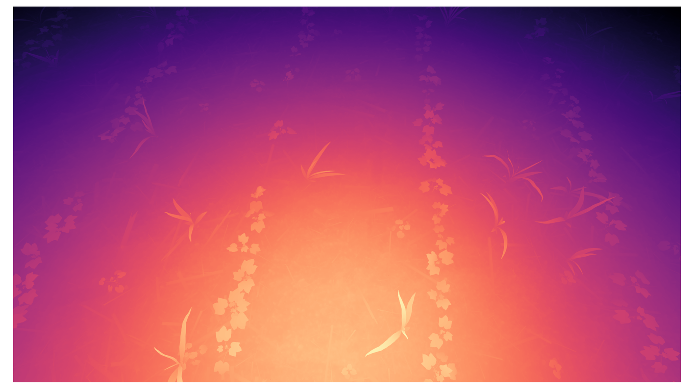
Includes diversity in:
- crop and weed growth stages
- soil color and moisture content
- crop residue type & distribution
- lighting conditions & time of day
- camera angle w.r.t to crop row
- crop and weed arrangement

Semantic masks



Instance masks

Depth masks



Domain adaptation

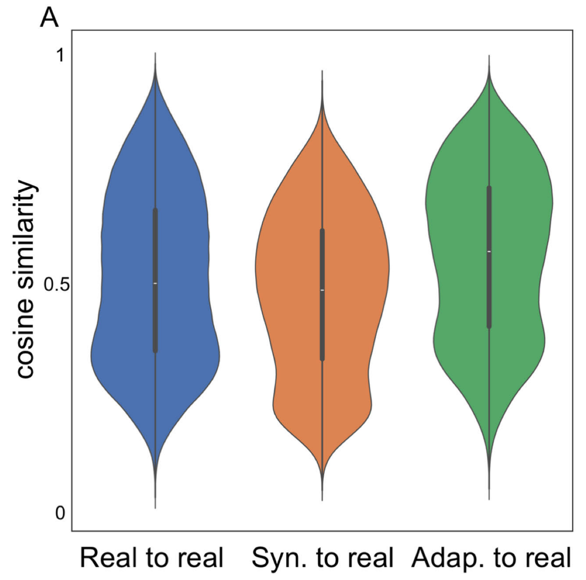
We used a GAN-based Contrastive Unpaired Translation (CUT) model* to translate from a synthetic domain to a real one. The model was trained on 1000 synthetic and real images.

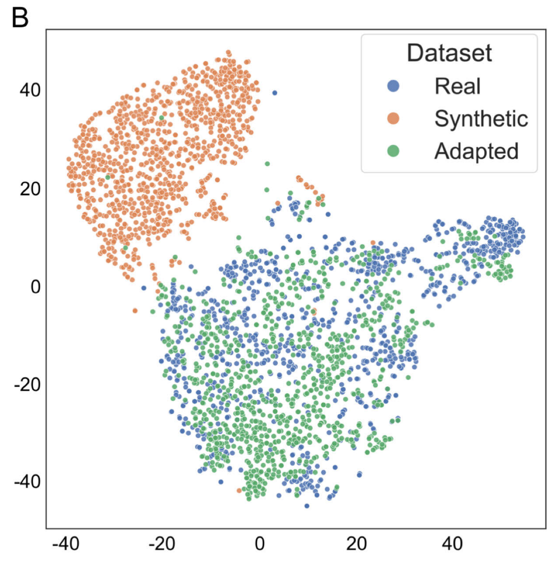
We compared images using feature vectors from a pre-trained, headless ResNet-50 network:
A) Cosine similarity
B) t-SNE visualization
C) Inception distance:

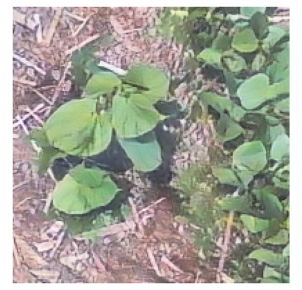
	Synthetic	Adapted
Fréchet ID	126.01	59.05
Kernel ID	0.0887	0.0188

Real Image

Domain-adapted Image







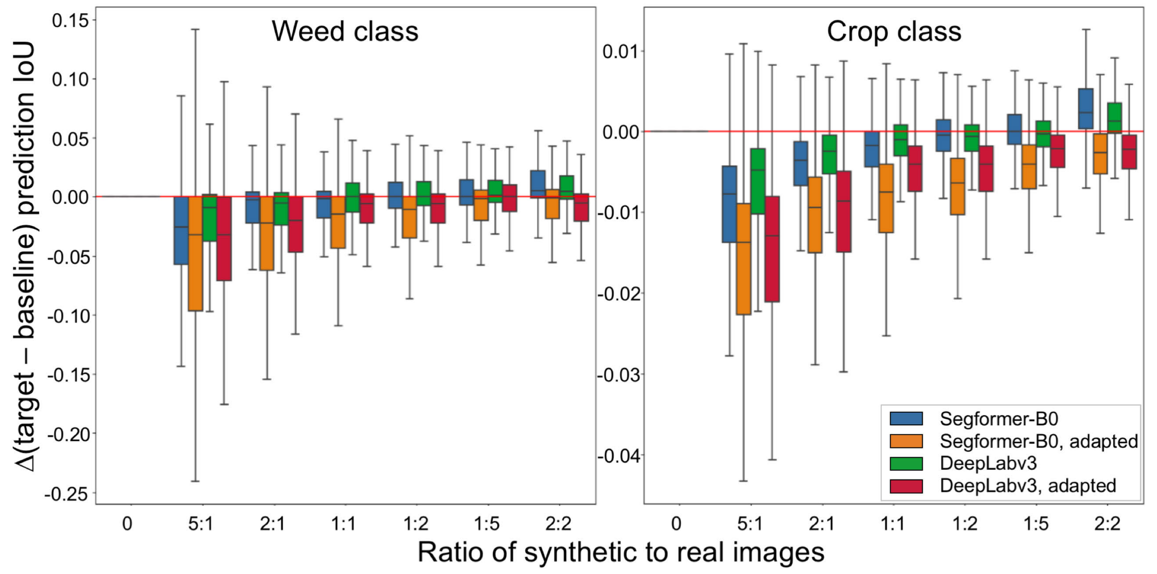
Close-up of domain-adapted image

*Taesung Park, Alexei A. Efros, Richard Zhang, and Jun-Yan Zhu. Contrastive learning for unpaired image-to-image translation. In *European Conference on Computer Vision*, 2020.

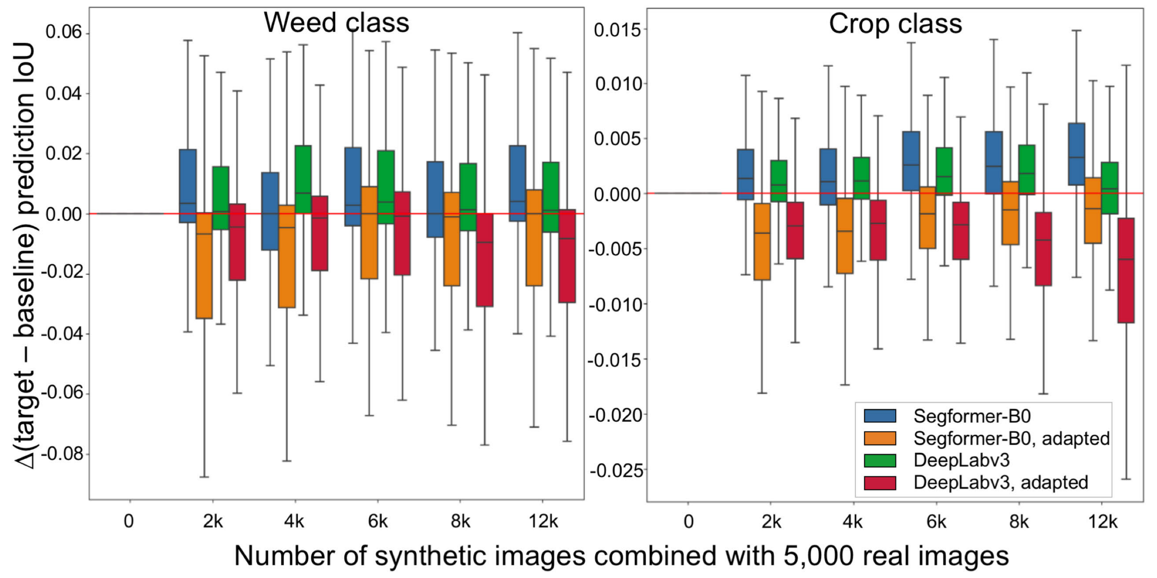
Evaluation

We fine-tuned SegFormer and DeepLabv3 models on a semantic segmentation task using real and synthetic images, and tested on a hold-out dataset of 3,981 real images.

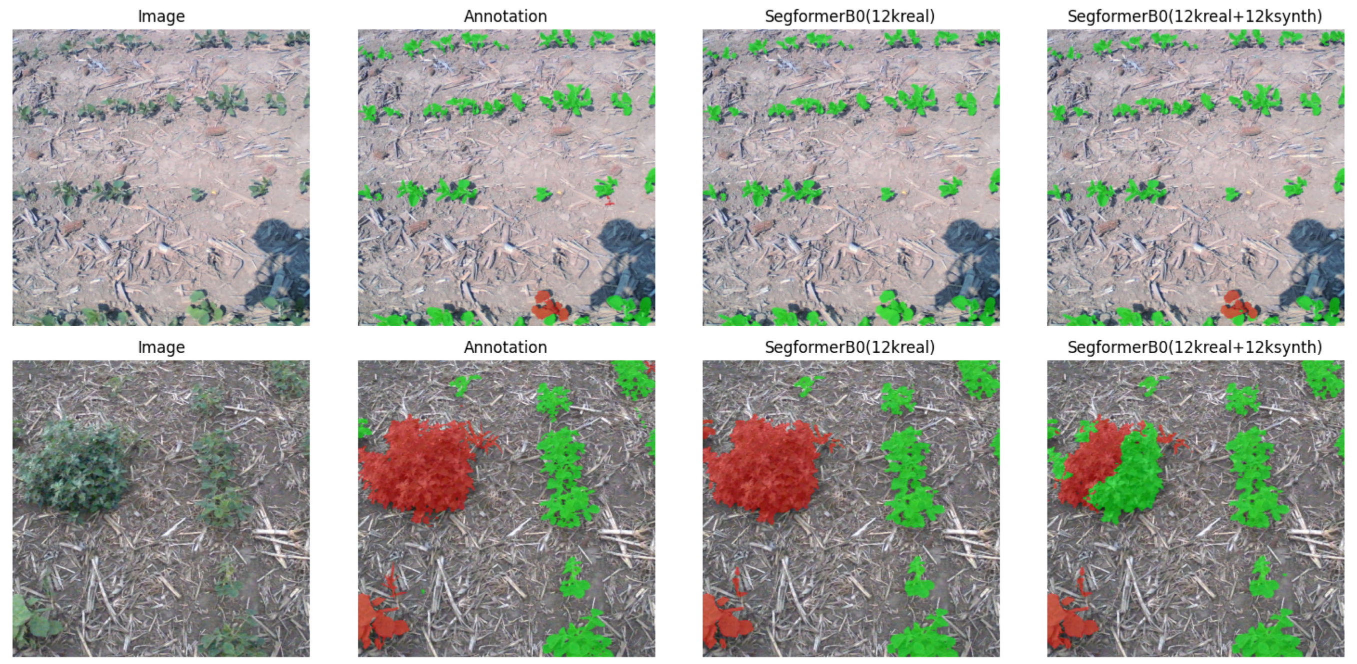
1) Difference in target and baseline IoU values for weed and crop classes with varying ratios of synthetic and real images. Combining synthetic and real images boosts performance.



2) Constant number of real images with varying number of synthetic images. Adding more synthetic images shows a small linear increase in performance for both classes and models.



3) Example images from a real dataset, with manually-annotated labels compared to two Segformer-based predictions: trained on 12k real images (third column), and 12k real + 12k synthetic images (last column).



Conclusions

1) Procedurally generated synthetic images are an effective data augmentation strategy, giving fine control of complex scenes.
2) Combining real and synthetic images improves model performance, including better model generalization and resiliency.
3) Synthetic-to-real domain adaptation using a CUT model did not improve performance on our segmentation task.

Train on soybean fields, test on cotton fields

syn : real ratio	mlou – Weed SegF	mlou – Weed DeepL	mIoU – Crop SegF	mIoU – Crop DeepL
12k : 0	0.0427↓	0.0359↓	0.4626↓	0.2904↓
10k : 2k	0.1203↓	0.1172↓	0.7232↑	0.6355↑
8k : 4k	0.1439↑	0.1348↓	0.7256↑	0.6718↑
6k : 6k	0.1446↑	0.1370↑	0.7214↑	0.6581↑
4k : 8k	0.1549↑	0.1409↑	0.7285↑	0.6559↑
2k : 10k	0.1553↑	0.1430↑	0.7155↓	0.6330↑
12k : 12k	**0.1667**↑	**0.1411**↑	**0.7268**↑	**0.6412**↑
0 : 12k	0.1429-	0.1353-	0.7173-	0.6289-