



Human-Centric Computer Vision with Synthetic Data

Alex Thaman

Chief Software Architect, Computer Vision
Unity Software

Unity (U)

We believe the world is a better place with more creators in it

AI @ Unity

Enable customers of all sizes to unlock the power of AI with Unity



Challenges with Image Datasets

Bias in Image Datasets of Humans



Datasets inherit the biases of the data collection strategy.

These biases are **infrequently measured**, as doing so requires extensive labeling.

Bias is **difficult to control** due to needing careful targeted data collection.



Bias in Facial Recognition Technology

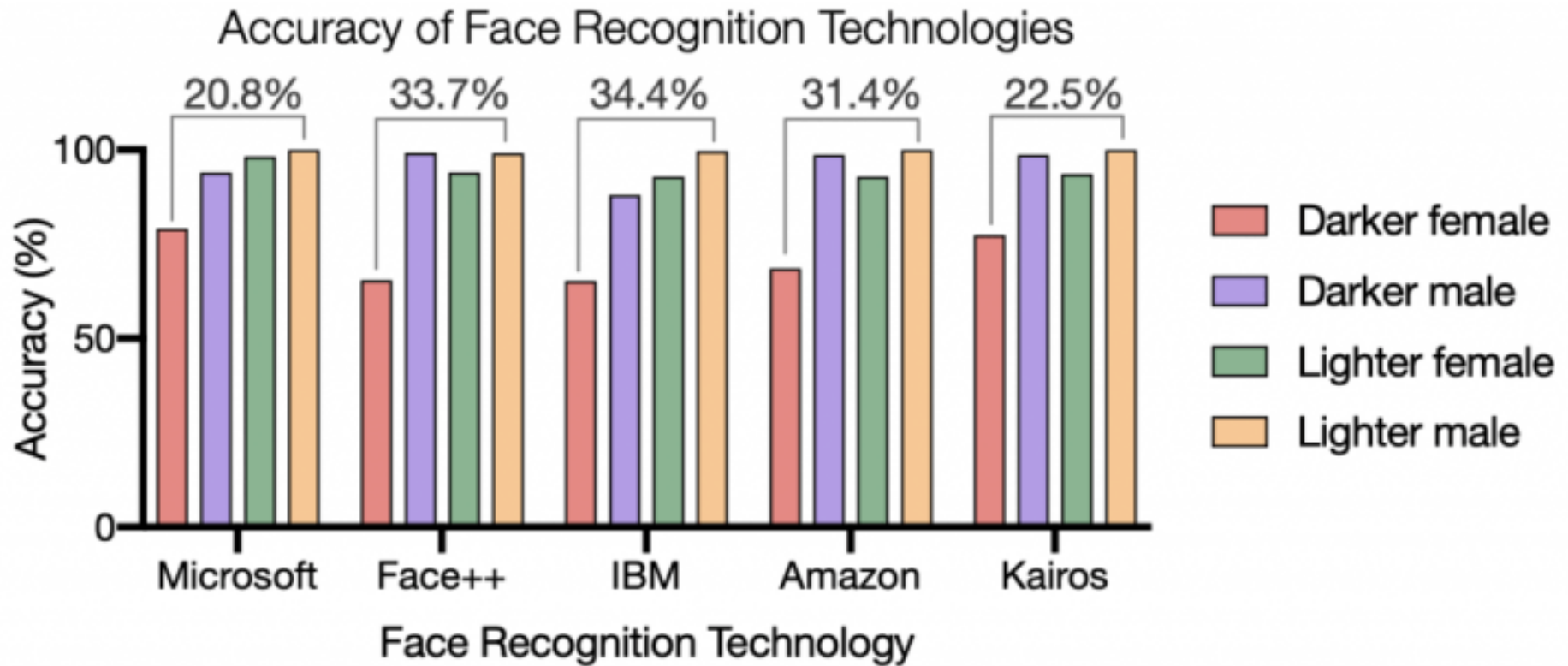
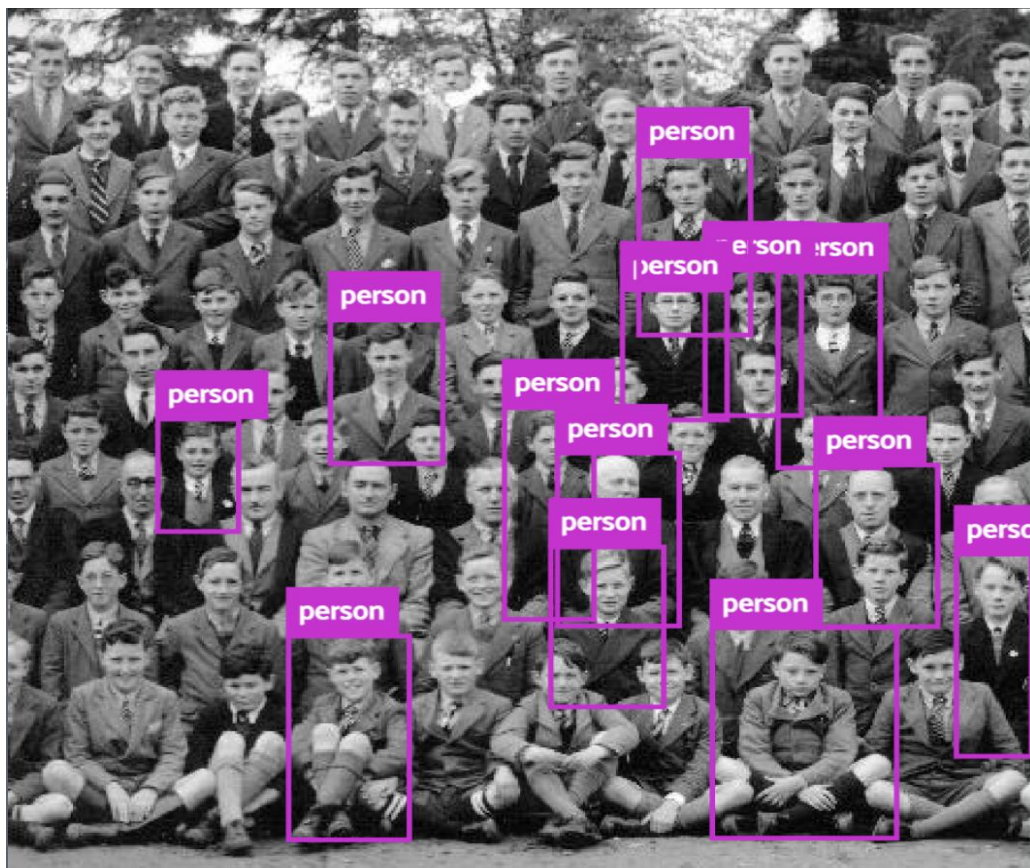


Image courtesy of sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology

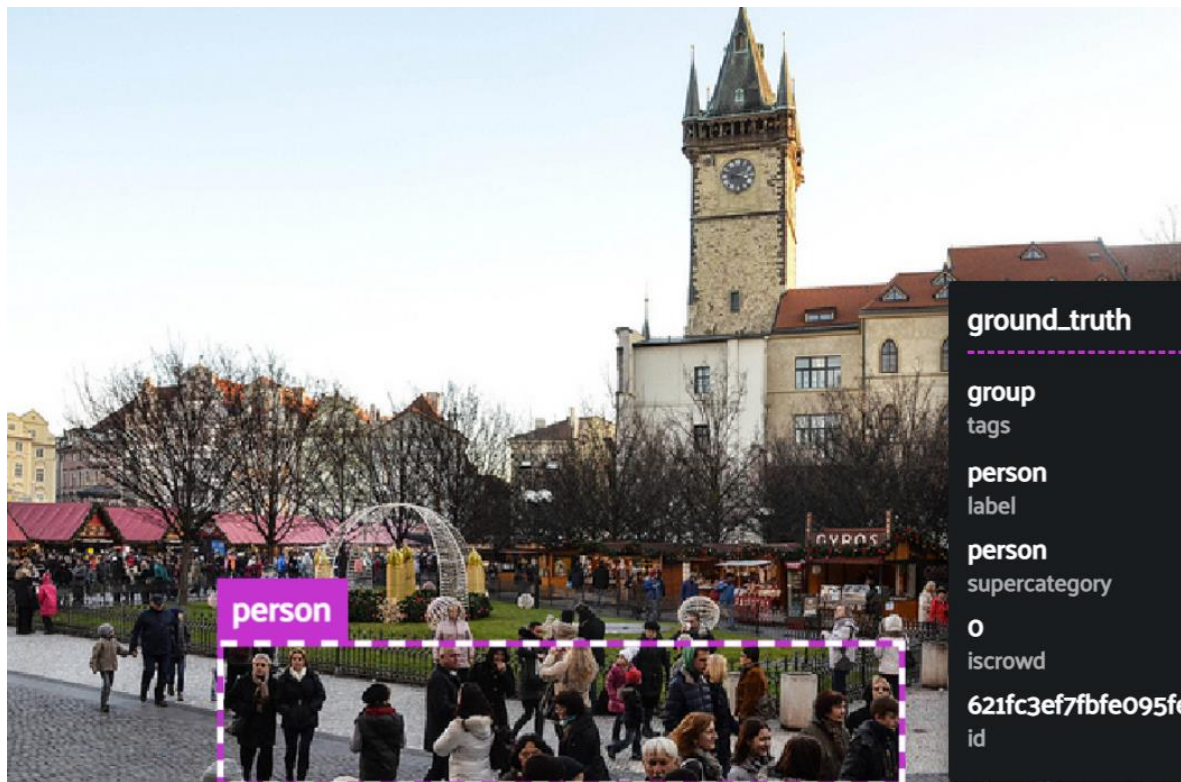
Missing Labels



Label Errors in Image Datasets of Humans



Incorrect Labels



Ambiguous/Inconsistent Labels



Measure and Control Datasets with Unity Computer Vision

- Unity is a real-time 3D game engine
 - Scene construction
 - Behavior authoring
 - Physics
 - Rendering
- Unity Editor for building games or simulators
 - Compatible with a wide range of industry-standard content creation tools (Blender, Maya, Houdini, Substance 3D Designer)
 - **A powerful tool for creating synthetic data**

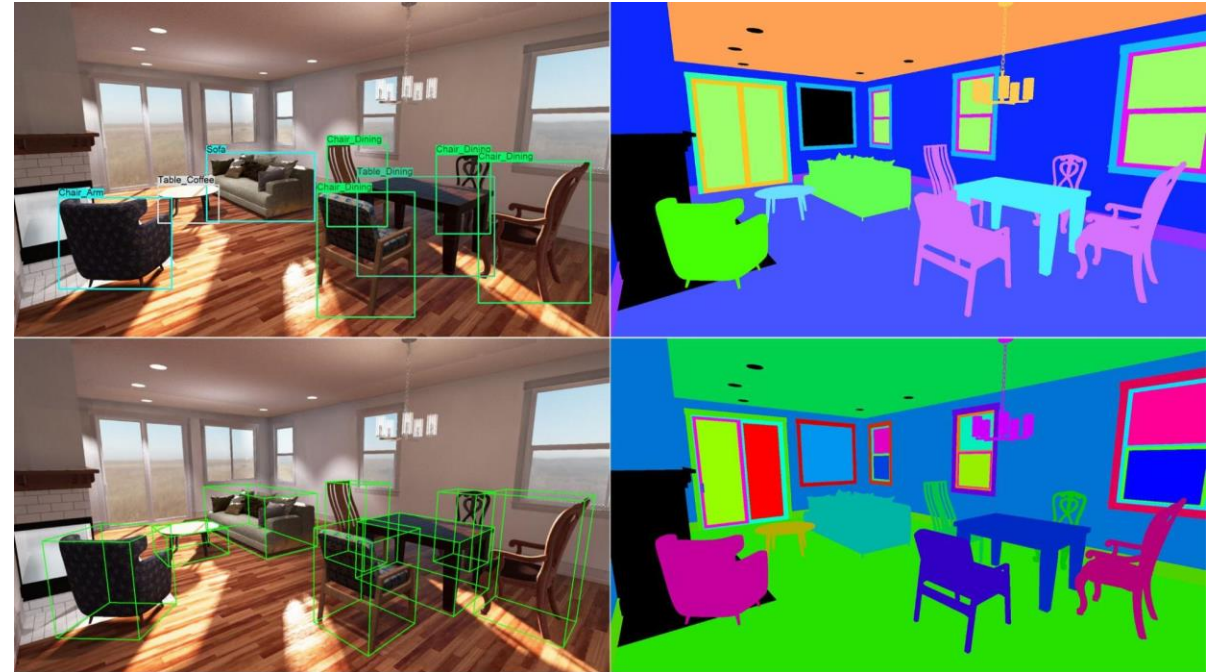


LGSVL Simulator, made with Unity (<https://www.svl simulator.com/use-cases/#automotive>)

Benefits of Synthetic Data



- Some real-world data is hard to collect and label
 - Unsafe
 - Rare
 - Privacy-protected
 - Labeling errors/inconsistencies
 - Biased towards collection conditions
- **Synthetic data** is:
 - Perfectly labeled
 - Perfectly controlled and safe
 - Privacy compliant

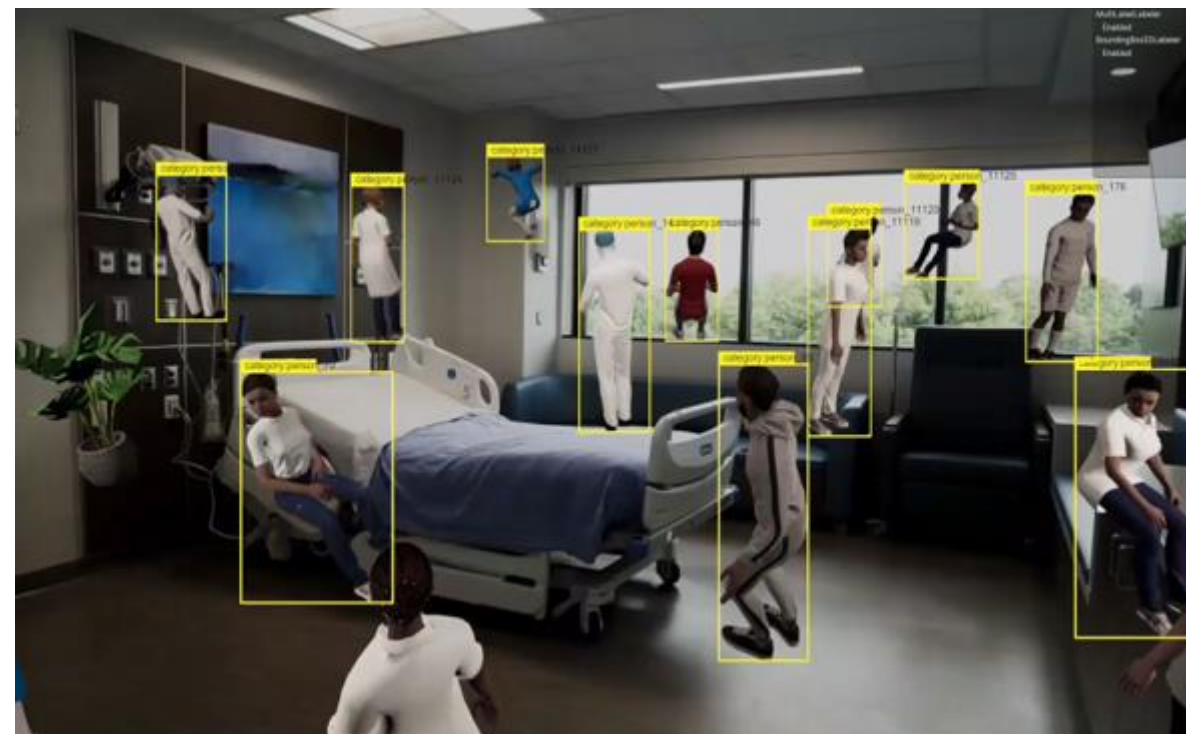


Challenge: Limited, biased image data leading to poor application performance in their ML model

Solution: Build synthetic datasets of room backgrounds with doctors/patients randomized in the scene using Unity Perception Package

Results:

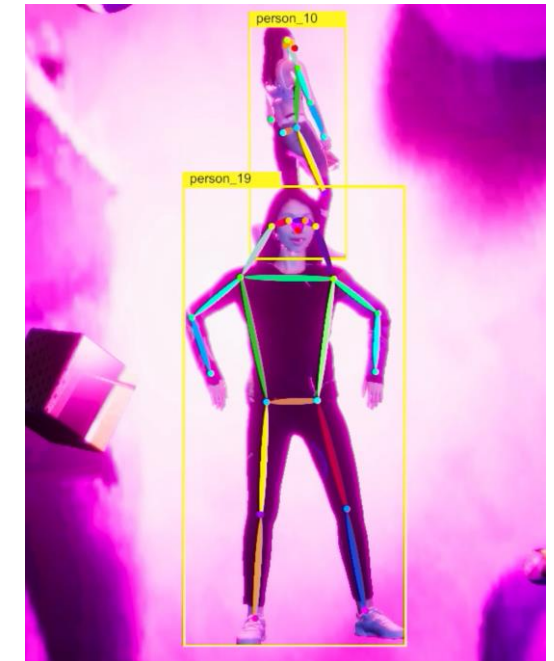
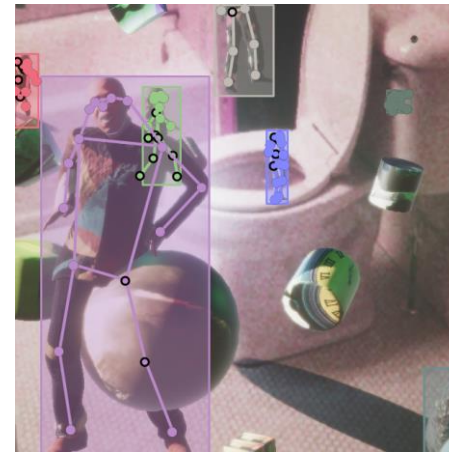
- 17% improvement in core KPIs
- Saved \$20K–\$40K in labeling costs
- Hours, instead of weeks, to create a **balanced** dataset



PeopleSansPeople

Benchmark for
Human-Centric Computer Vision

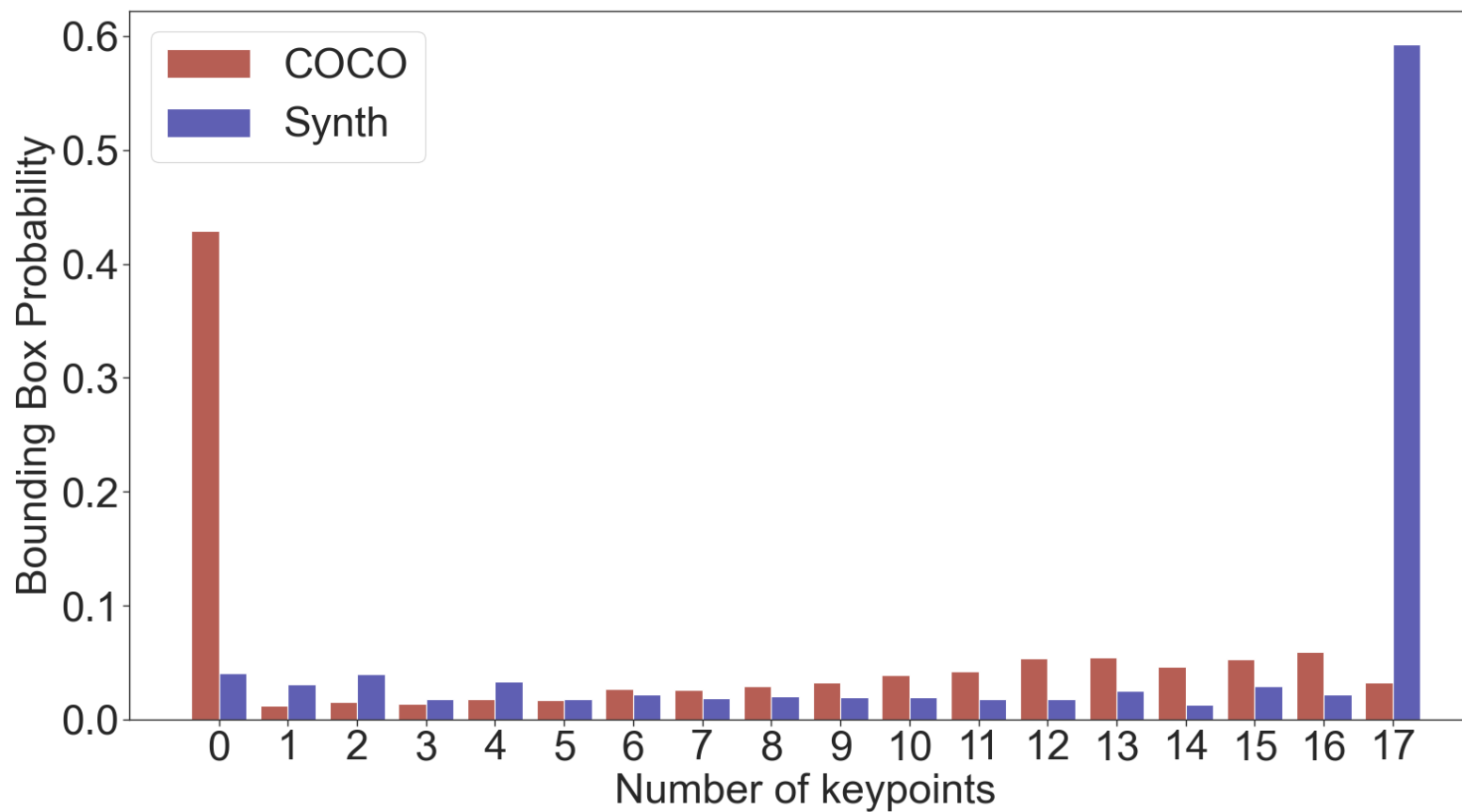
- **PeopleSansPeople**
 - Research benchmark
 - 29 scanned humans from RenderPeople
 - Random poses and clothing color
 - Random backgrounds, occluders, and distractors
 - Random lighting and post-processing



Unity Human-Centric Computer Vision



Perfect labeling – synthetic data provides all keypoints when they are visible



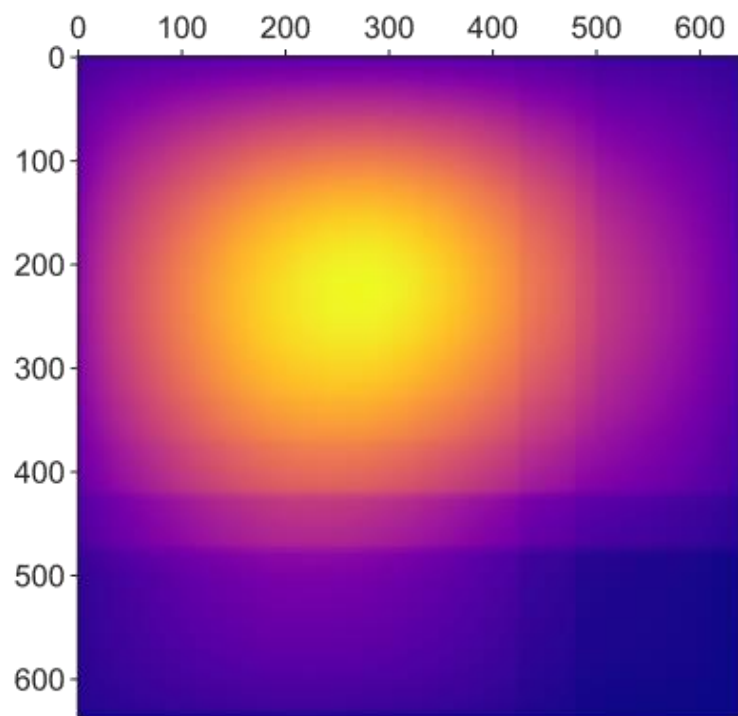
PeopleSansPeople: A Synthetic Data Generator for Human-Centric Computer Vision ([Arxiv 2112.09290v1](https://arxiv.org/abs/2112.09290v1))



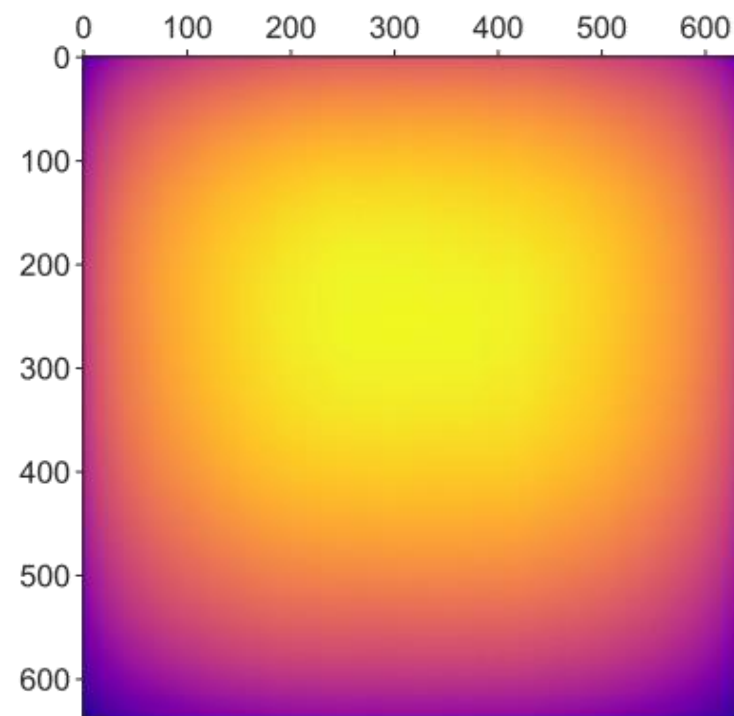
Unity Human-Centric Computer Vision



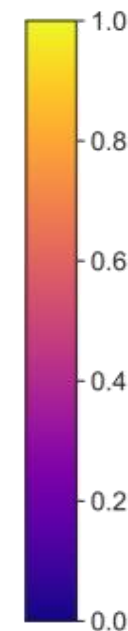
Measuring and controlling distributions – object location within image can be evenly distributed



(a) COCO



(b) Synthetic



PeopleSansPeople: A Synthetic Data Generator for Human-Centric Computer Vision ([Arxiv 2112.09290v1](https://arxiv.org/abs/2112.09290v1))

Pretraining with large amounts (~500K) of synthetic, unstructured images consistently improves all baselines.

		bbox AP (person val2017)				keypoint AP (person val2017)				
COCO	scratch	w/ ImageNet	w/ Synth	$\Delta/scratch$	$\Delta/ImageNet$	scratch	w/ ImageNet	w/ Synth	$\Delta/scratch$	$\Delta/ImageNet$
6411	37.82	42.53	48.97	+11.15	+6.44	39.48	45.99	55.21	+15.73	+9.22
32057	52.15	52.75	54.93	+2.78	+2.18	58.68	60.28	63.38	+4.70	+3.10

		keypoint AP (test-dev2017)			
COCO	scratch	w/ ImageNet	w/ Synth	$\Delta/scratch$	$\Delta/ImageNet$
6411	37.30	44.20	52.70	+15.40	+8.50
32057	55.80	57.50	60.37	+4.57	+2.87

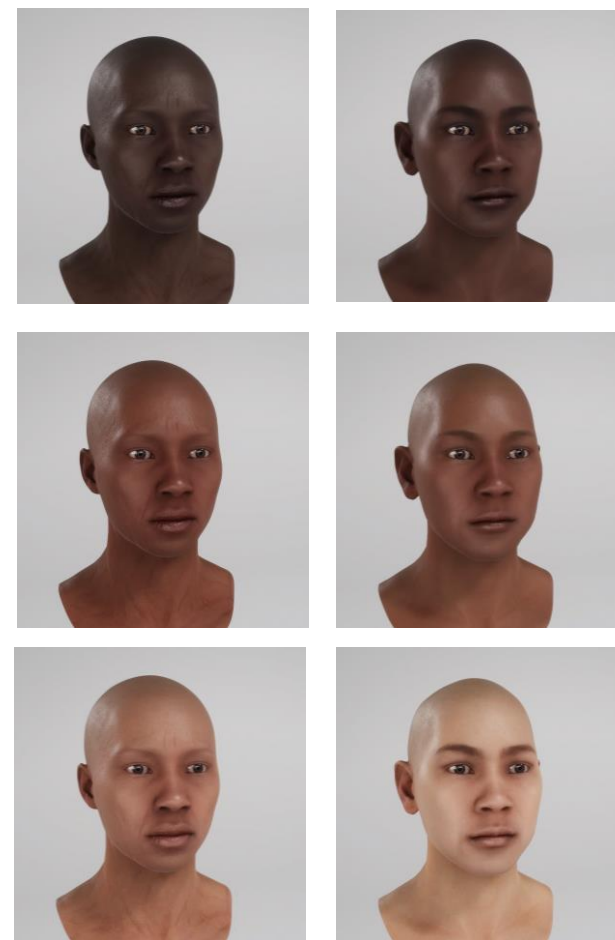
Unity Human Datasets

- Controlling human diversity in image datasets is critical to avoid bias and spurious correlations that affect model performance
- Introducing **Human Datasets with Unity Computer Vision**
 - Fully anonymized human images – our images of humans do not represent any actual humans that exist in the world
 - Height, weight, clothing, hair, accessories, and pose are independently sampled, or sampling can be tailored to use case
 - Perfect labeling of body keypoints, 2D/3D bounding box, segmentation variants, and depth
 - Additional human metadata for diversity characteristics (ethnicity, age, etc.)

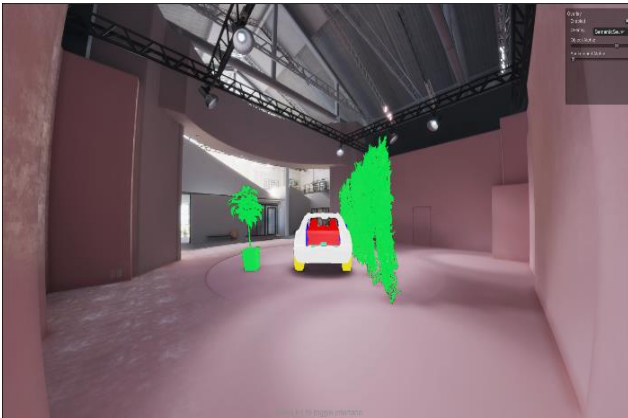
Unity Human Datasets



Unity Human Datasets



- Additional features available for all of our datasets:
 - Configurable RGBD sensors, including fisheye and other distortions
 - Post-processing effects randomization – hue, blur, focus, contrast, saturation
 - Environmental randomization – scene layout, lighting, HDRI backgrounds
 - Scaling up data generation in Unity Computer Vision Datasets service



- The Unity AI group is bringing **content creators** together with **AI creators** into a collaborative, accessible ecosystem
- Unity's investment in real-time 3D tools and services will remove boundaries and democratize AI development, starting with access to data
- Recent acquisitions of Weta and Ziva increase the Unity tools suite for creating digital humans, animals, and other creatures



wētā
DIGITAL



Call to Action



- Read our blog and research publication for PeopleSansPeople, and visit the related GitHub page (see Resources slide)
- Join our Computer Vision Customer Program for access to new Unity tools for building your own human-centric computer vision datasets: computer-vision@unity3d.com
- Interested in joining our team? We are hiring research scientists, ML engineers, software engineers, and artists. Reach out to me directly.

Call to Action



Contact us about custom human-centric computer vision datasets for your use case



Additional Resources



Product page

<https://unity.com/products/computer-vision>

PeopleSansPeople blog

<https://blog.unity.com/technology/human-centric-computer-vision-with-unity-synthetic-data>

PeopleSansPeople GitHub

<https://github.com/Unity-Technologies/PeopleSansPeople>

Ouva webinar

<https://resources.unity.com/ai-ml/ouva-synthetic-training-data-webinar>

My contact info

Alex.Thaman@Unity3D.com

<https://www.linkedin.com/in/alex-thaman-93436659/>



Thank You.

