

# Enhancing Sustainable Crop Production with Machine Learning, Synthetic Data, and Digital Twin for Strawberry

**Daeun “Dana” Choi**

Assistant Professor  
Smart Agriculture Laboratory  
Department of Agricultural and Biological Engineering  
Gulf Coast Research and Education Center  
University of Florida

## Outline



**Optimizing Machine Vision & Robotics in Agriculture Through Virtual Environments**

### Introduction

- Challenges in Ag Automation
- What is Synthetic Data?

### Digital Twins for Strawberry Farming

- How to Create Realistic Synthetic Vegetative Scene
- Creating a Scene: Hardware Control

### Similarity between real and synthetic data

- Fruit Detection
- Fruit Sizing
- Lidar Measurement

### Conclusion & Discussion



**Machine  
learning and  
field robotics for  
precision  
agriculture**

# **Smart Agriculture Lab**

Where Agriculture & Technology Meet to Build Future

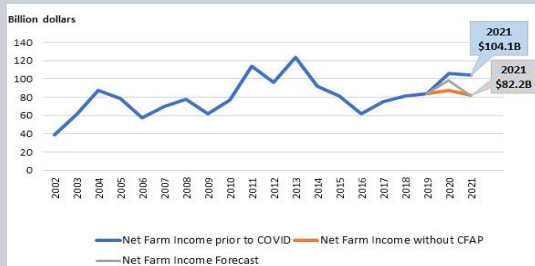


# **Introduction**



# Challenges in Agriculture

1. Decrease in agricultural land
2. Labor scarcity in rural areas
3. Price and availability of energy
4. Climate change



Graph from USDA, America's Farmers: Resilient Throughout the COVID Pandemic, Oct 13, 2020

AP AP News

Farm labor shortage nothing new, getting worse, farmers say  
And the struggles have become that much harder since labor shortages have hit other industries as pandemic shutdowns ease. ADVERTISEMENT. Those ...  
2 weeks ago



WKRG

Weather, labor shortages and high demand causing food inflation

According to the U.S. Department of Agriculture, grain that feeds ... weather conditions and labor shortages from the pandemic is the ...  
15 hours ago



USA Today

Finding workers was already hard for the ag industry. Now, it's even worse, farmers say

And the struggles have become that much harder since labor shortages have hit other industries as pandemic shutdowns ease.  
7 weeks ago



Economic Perspective

... after-effects just beginning | Agriculture ...

... has been recovering since third-quarter 2020. ... First, labor ... a drag on the economic recovery ...



# Precision Agriculture

## Optimization of Farming Input

- Efficient farming by site-specific crop management
- Increase crop yields
- Maximize profit for growers
- Reduce environmental impact

Technologies  
for Enhanced Precision Agriculture

- **AI and Machine Learning:** data analytics and forecasting
- **Robotics:** Automation with increased efficiency
- **Digital Twin:** digital representation of your farm



Treatment required areas

Citrus grove in Lake Wales, Florida

# Challenges in Ag Automation & Robotics



## Time-Consuming

Testing is limited to few months



## Special Expertise

Installation  
Maintenance



## Costly

Data collection  
Travel to farm site



## Risk

Bad data  
Crop is lost due to disease or frost



**How to reduce the turnaround time for autonomous systems?**

# Synthetic Data

Definition:

- Data generated artificially, not from real-world observations

Purpose:

- Can replicate the characteristics of real data
- Useful for enhancing datasets, privacy, and testing

Generation Techniques:

- Statistical methods (e.g., bootstrapping, synthetic regression)
- Generative models (e.g., GANs)
- Simulations & 3D graphics

**Digital Twin:** a dynamic, virtual representation of a physical object or system, allowing for real-time “artificial” monitoring or simulation

Shape your land into a thriving farm

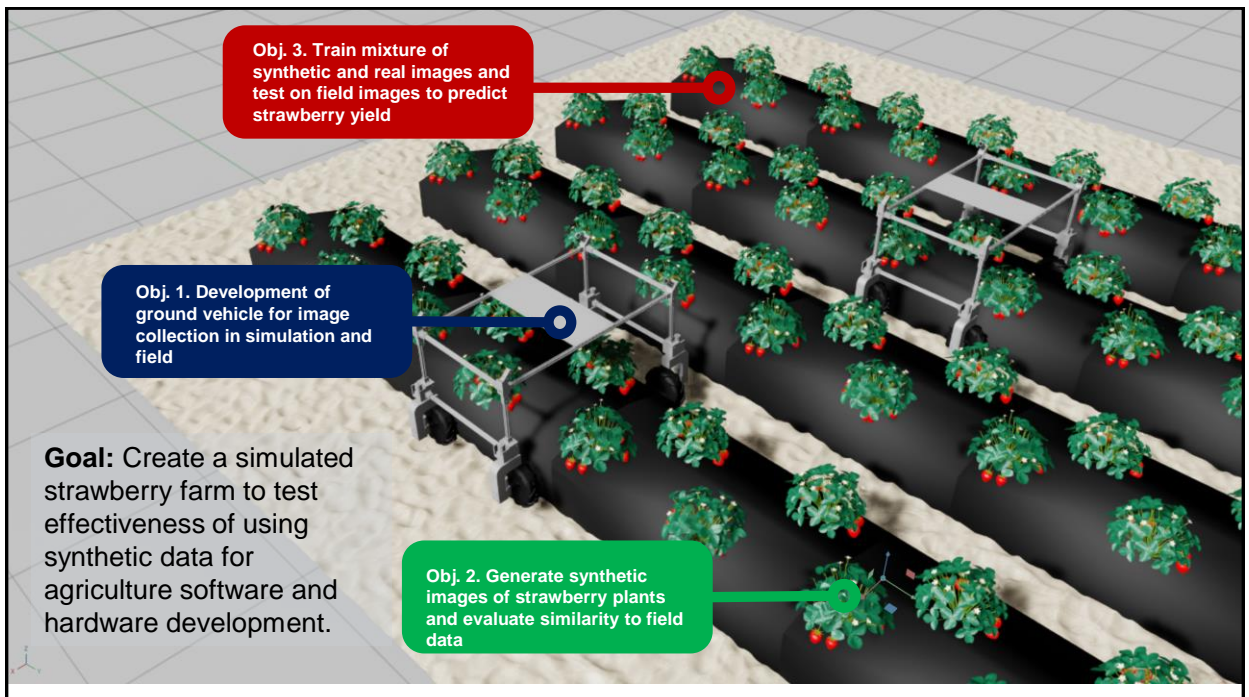
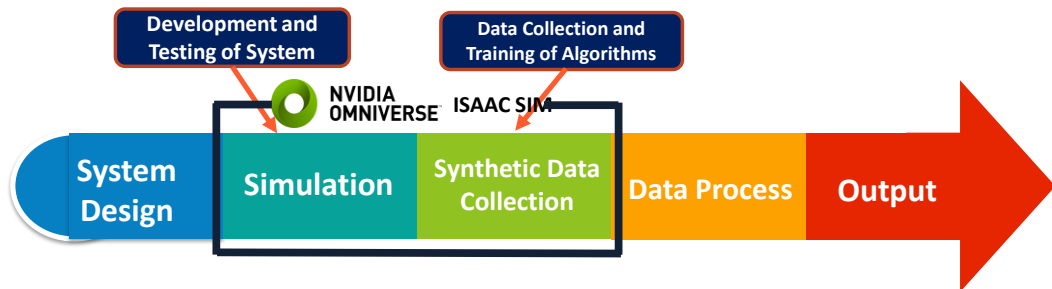
## Exploring the Virtual Strawberry Farm



# Benefits of Digital Twin

Speed up the development process of robots and AI

- Data augmentation for limited datasets.
- Training robust machine learning models.
- Testing new agricultural tools without risk.
  - Ensuring farmers' data privacy.

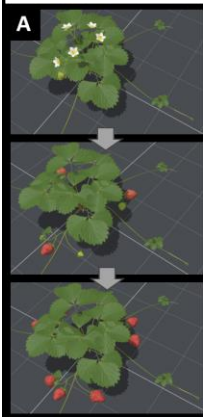






# Procedural Modeling

- Rules-based approach for modeling rather than manual design of components
- Flexibility: Easier to modify model components instead of recreating the entire model again
- Randomization: Can randomize model

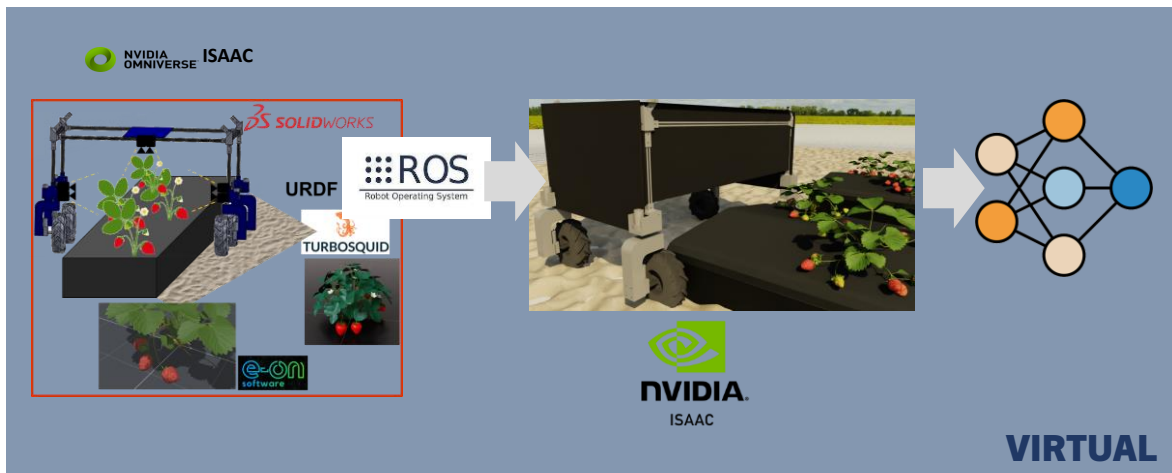


Clump radius	<input type="range"/>	0.584347
Clump angle	<input type="range"/>	0.634783
Leaf stalk length	<input type="range"/>	0.226086
Leaf stalk angle	<input type="range"/>	0.930435
Leaf stalk curving	<input type="range"/>	0.32826
Leaf scale	<input type="range"/>	0.417391
Leaf width	<input type="range"/>	0.617391
Leaf stalk axial rotation	<input type="range"/>	0.521739
Flower stem length	<input type="range"/>	0.730435
Flower stem angle	<input type="range"/>	0.391305
Sepal length	<input type="range"/>	1
Sepal width	<input type="range"/>	1
Petal length	<input type="range"/>	0.965217
Petal width	<input type="range"/>	0.660871
Fruit scale	<input type="range"/>	0.53913
Fruit length	<input type="range"/>	0
Fruit shape change	<input type="range"/>	0
Fruits warp intensity	<input type="range"/>	0.608696
Flower and fruit growth	<input type="range"/>	0.626087
▼ MOTHER PLANT DENSITY		
Plant unit (1 to 10)	<input type="range"/>	0.0173916
Leaf petiole (3 to 13 +/- 1)	<input type="range"/>	0.373913
Flower or fruit stem (1 to 5 +/- 1)	<input type="range"/>	0.66087
Fruit warped ratio	<input type="range"/>	0

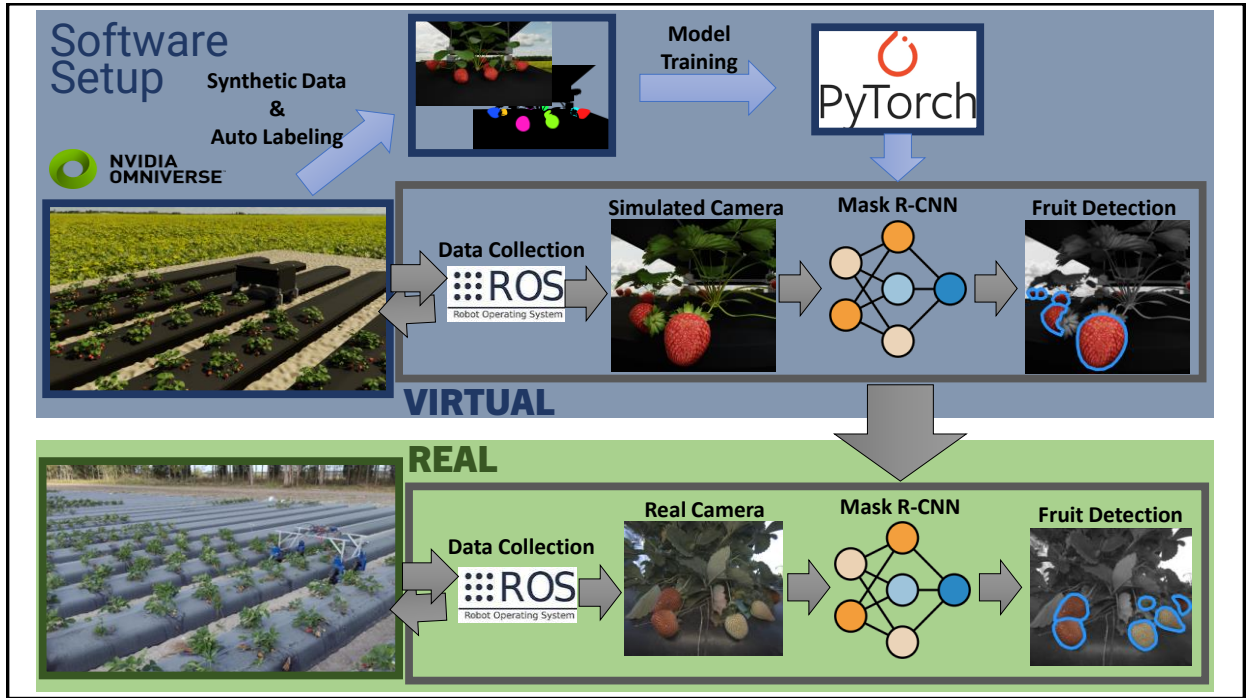
# Hardware Setup in the Field



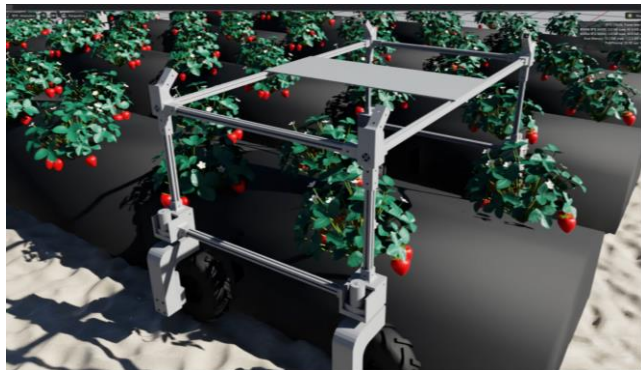
# Hardware Setup in Isaac Sim



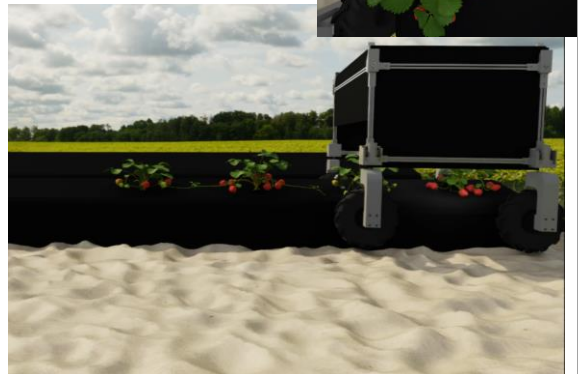




## Data Acquisition (Gen1 & Gen2)



Gen 1



Gen 2

Field

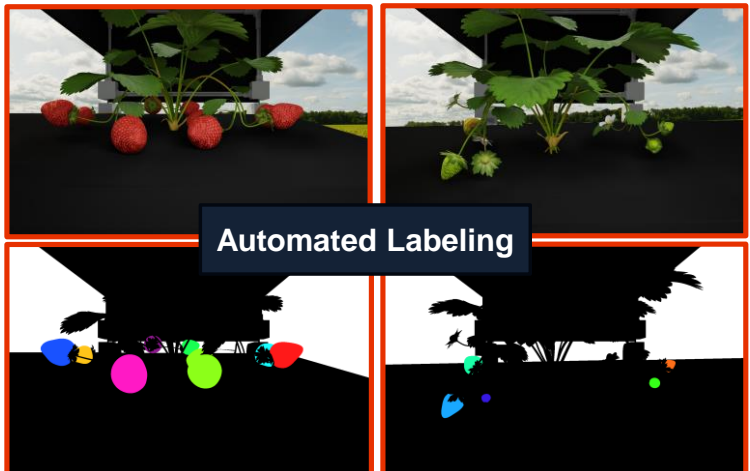


Simulation



## Camera Perspectives

## Automated Labeling

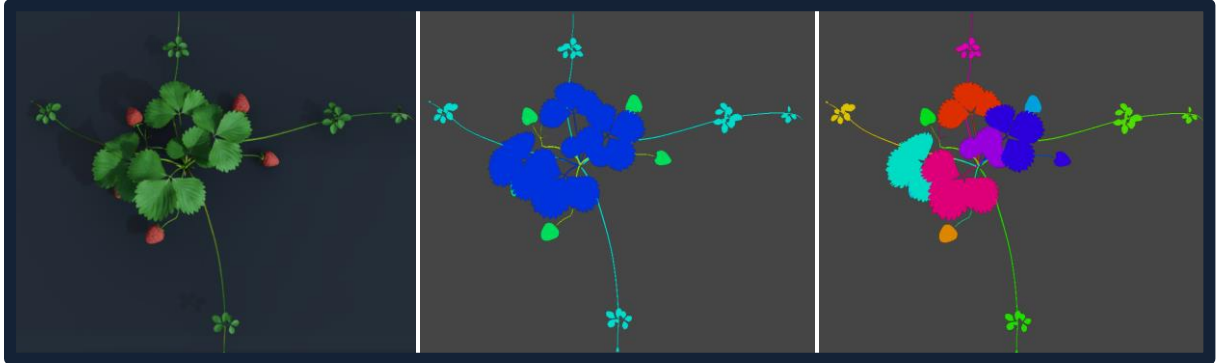


# Automated Labeling

RGB

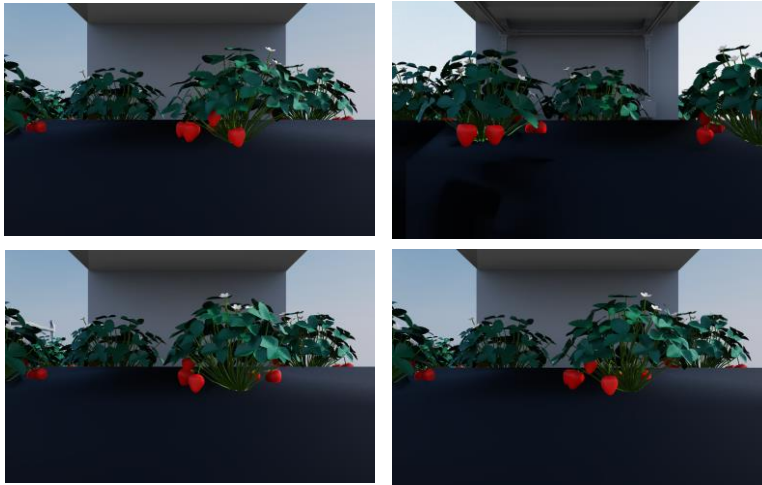
Semantic  
Segmentation

Instance  
Segmentation

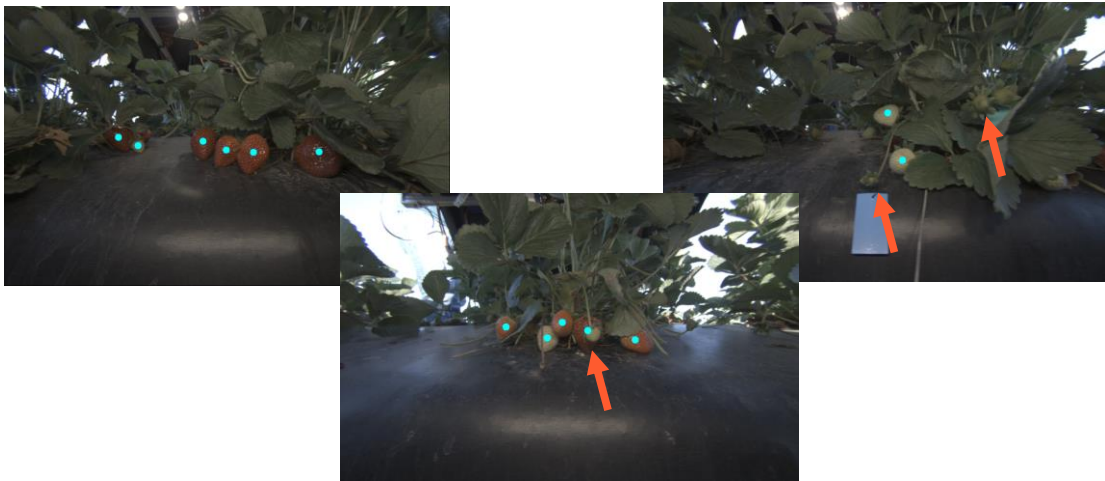




## Results: Sample Training Images - Gen 1



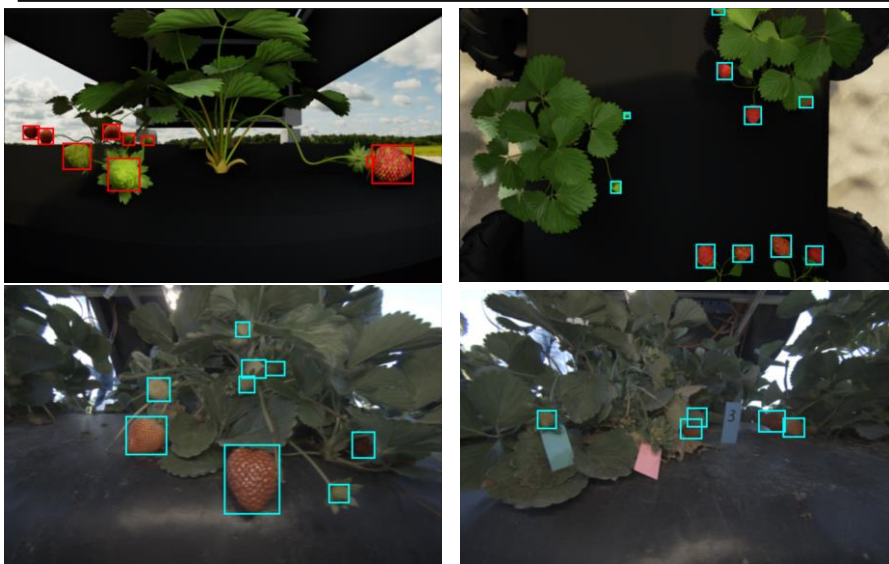
## Results: Fruit Detection on Field Images



## Sample Training Images - Gen 2





## Results: Fruit Detection on Field Images



# Results: Fruit Detection

Training using Synthetic Data only

	TP	FP	FN	Precision	Recall	F1-Score
 Synthetic Data Gen-1	56	2	86	0.96	0.39	0.56
 Synthetic Data Gen-2	127	17	15	0.88	0.89	0.89
Gen-1 + Gen 2	96	3	44	0.97	0.69	0.80
Real Images	129	2	12	0.98	0.91	0.95

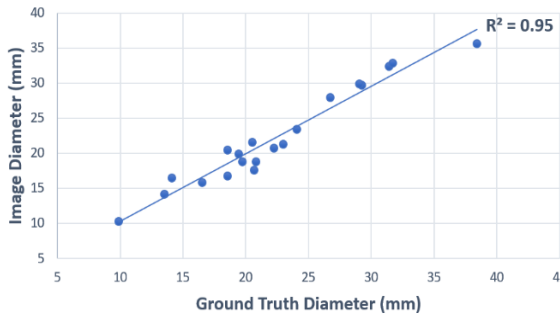
# Results: Fruit Detection

	Trial 1 (March 16)			
	Fruit Count	Precision	Recall	F1-Score
All Fruit	295	0.95	0.89	0.92
Red Fruit	167	0.95	0.99	0.97
White Fruit	61	1.00	0.87	0.93
Green Fruit	67	0.94	0.72	0.82
	Trial 2 (March 20)			
All Fruit	134	0.89	0.75	0.81
Red Fruit	39	0.87	0.95	0.91
White Fruit	52	0.98	0.81	0.89
Green Fruit	43	0.83	0.58	0.68



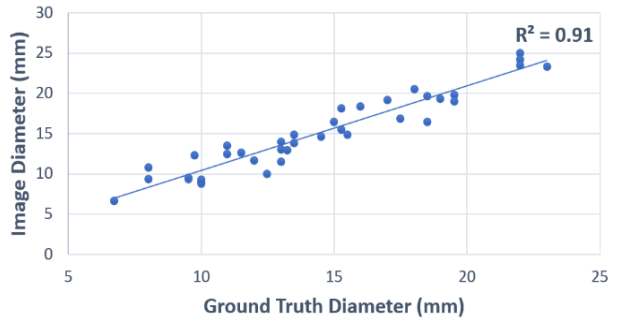
# Results: Fruit Sizing

Simulation Based Diameter Estimation



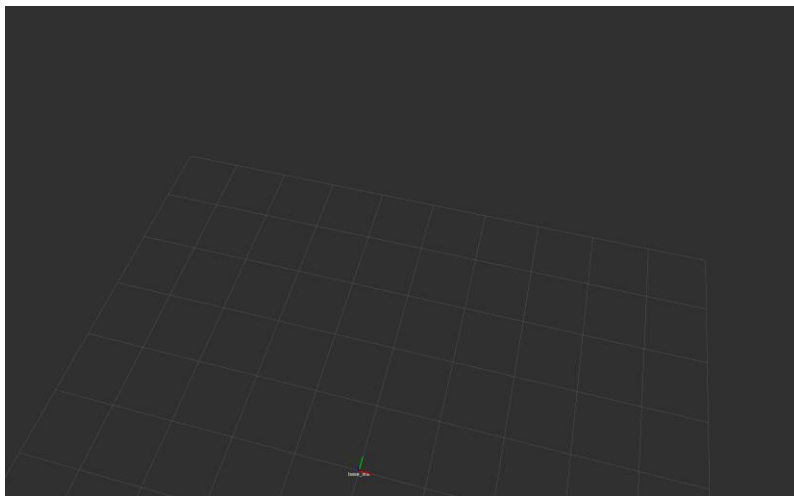
Average Diameter Error: 1.5 mm

Field Based Diameter Estimation

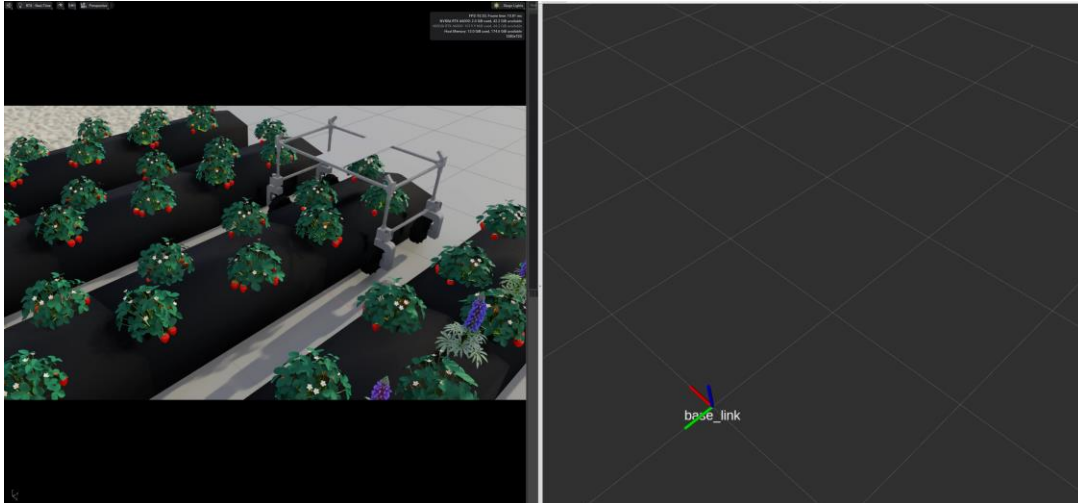


Average Diameter Error: 1.6 mm

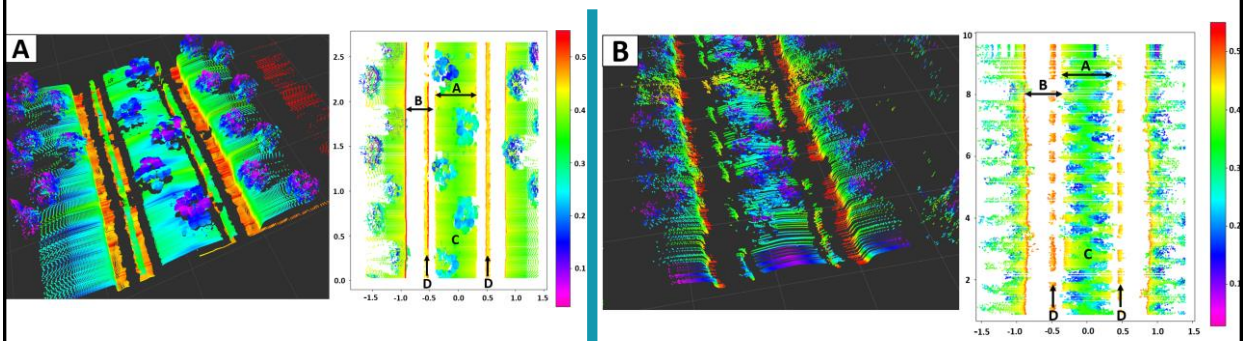
# Results: Lidar Data In Field



# Results: Lidar Data In Simulation



# Results: Lidar Data Comparison

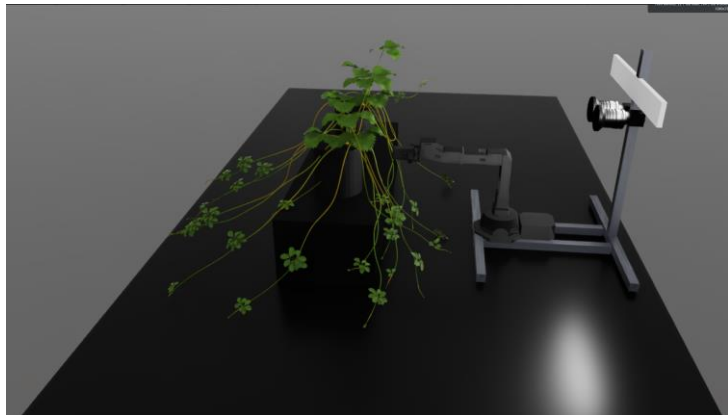


Label	Description	Simulated Lidar Sensor (m)	Real Lidar Sensor (m)	Ground Truth Measurement (m)
A	Bed Spacing	0.50	0.53	0.51
B	Bed Width	0.72	0.71	0.74
C	Bed Height	0.26	0.29	0.25

## More realistic strawberry Plant - Runner

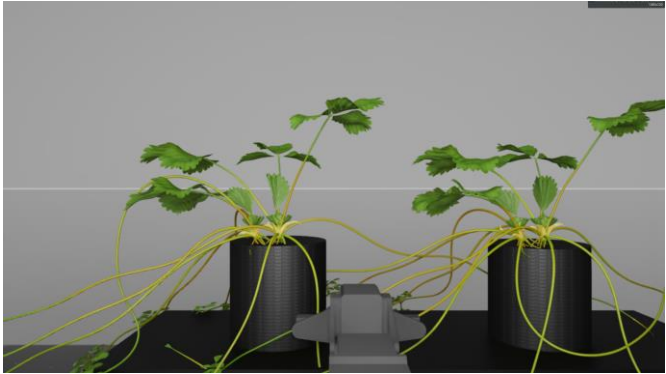


## Digital Twin for Runner Cutting



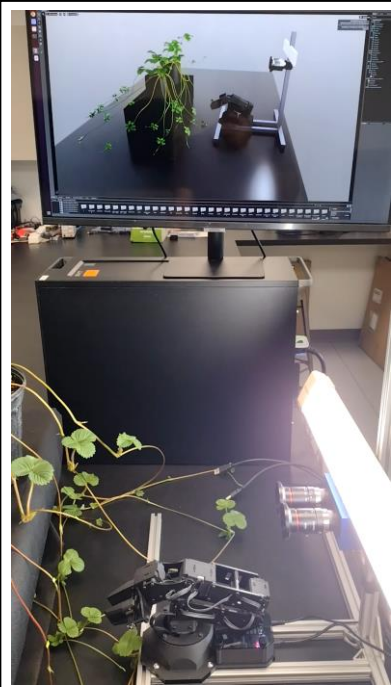


# Digital Twin for Runner Cutting



# Digital Twin for Runner Cutting

Bidirectional communication between ROS and the hardware, as well as its reflection in the virtual environment of Isaac Sim



# Robotics Simulation for Runner Detection



## Summary

**Turnaround time from concept to product in the AgTech industry is costly and time-consuming.**

- There is a limited window to test prototypes each year, which adds to the challenges.
- The risk of crop limitations due to disease or weather further complicates the development process.
- Labor costs associated with collecting data and obtaining reliable ground truth are significant.

**Leveraging synthetic agricultural data and better simulation tools can help mitigate risks.**

- Utilizing synthetic data allows for more frequent and extensive testing, reducing reliance on limited real-world opportunities.
- Improved simulation tools provide a more efficient and cost-effective way to evaluate product performance.

# Acknowledgement

---



Omeed Mirbod  
PhD Student



Uchechukwu Ilodibe  
PhD Student



Namrata Dutt  
PhD Student

**Thank You!**



ANY QUESTIONS?  
YOU CAN FIND ME AT:  
DANA.CHOI@UFL.EDU