CSE 244A Final Project “Image Classification with Semi-Supervised Learning”

Nick Wang, Akashleena Sarkar

Agenda:

- Problem Statement
- ResNet 50 Baseline Method and Results
- Experimenting with DeIT and Results
- Experimenting with DaViT and Results
- Future Work
- Work Division

What is this about?

Problem Statement:

- Need for effective image classification with limited labeled data.

Goal:

- To evaluate the performance of AI architectures and fine tune parameters t maximize accuracy

Approach Overview:

- Train multiple models using both labeled and unlabeled data.
- Compare accuracy and performance metrics.

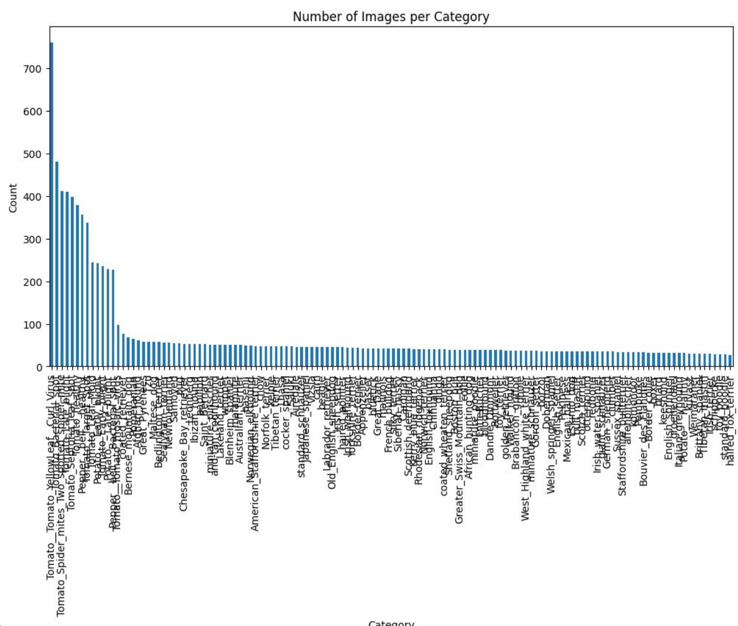
Project Goals and Objectives

Dataset Description:

- 15 Plant Categories, 120 Dog Categories.
- Total labeled and unlabeled images.

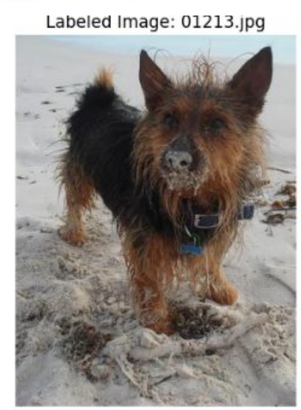
Challenges:

- Imbalanced categories.
- Noise in unlabeled data.



Number of training images: 7883
Number of validation images: 1971

Number of unlabeled images: 22995



ResNet 50 - Data Preprocessing and Data Augmentation

Preprocessing Steps:

- Images resized to **224x224 pixels**.
- Normalization using ImageNet statistics:
  - Mean: [0.485, 0.456, 0.406]
  - Std: [0.229, 0.224, 0.225]
- Validation set transformations:
  - Resized to fixed size (224x224).
  - Normalized without additional augmentations.

Data Augmentation for Training:

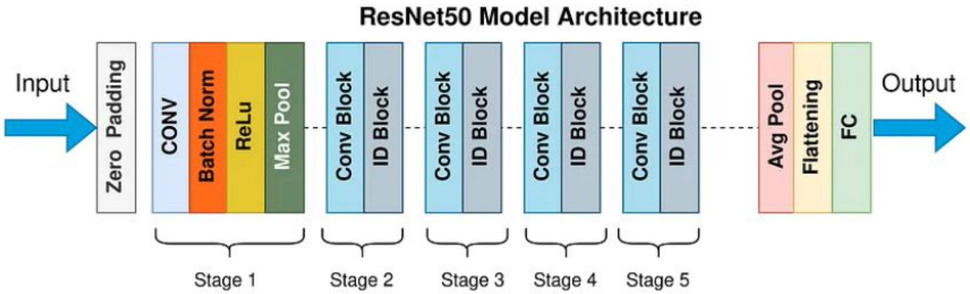
- **Random Resized Crop:** Simulates zoom variations.
- **Random Horizontal Flip:** Increases robustness to mirrored images.
- **Color Jitter:** Adjusts brightness, contrast, saturation, and hue.
- **Random Rotation:** Rotates images by ±15° for orientation diversity.

ResNet 50 - Model Architecture and Fine-Tuning

Model Used: Pre-trained ResNet50 (ImageNet weights)

Layer Adjustments:

- Final fully connected (fc) layer replaced to match **134 classes**
- Layer 3 and layer 4 and **fc** were unfrozen for fine-tuning:
    - Higher-level features refined for task-specific knowledge.



Resnet-50 Model architecture

Why ResNet50?

- Robust architecture for image feature extraction.
- Transfer learning leverages pre-trained knowledge for faster convergence.

ResNet 50 - Training Setup

Training Parameters:

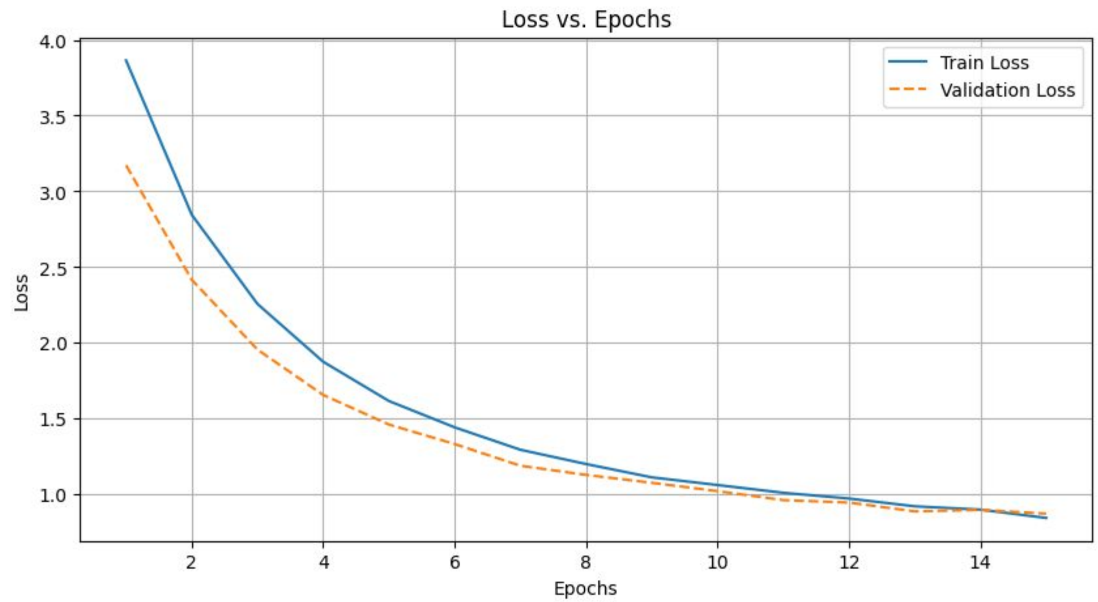
- **Number of Epochs:** 15
- **Batch Size:** 32
- **Loss Function:** CrossEntropyLoss
- **Optimizer:** AdamW with:
  - Learning rate: **1e-4** (fc), **1e-5** (layer4)
  - Weight decay: **1e-4**
- **Learning Rate Scheduler:** StepLR reduces learning rate by **0.1** every **5 epochs**.

Device: CUDA-enabled GPU for accelerated training.

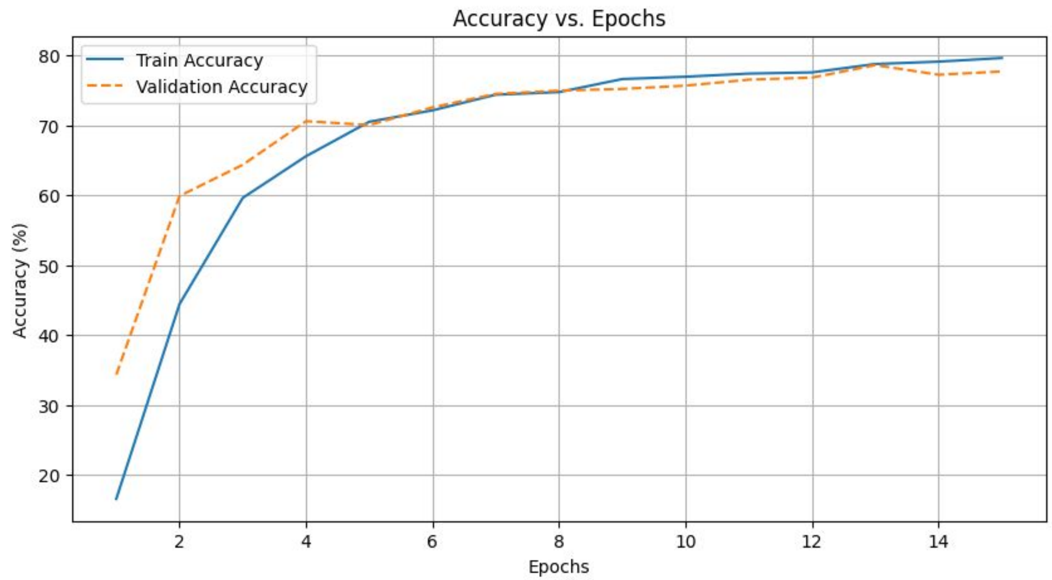
Data Loaders:

- Training batches: 247
- Validation batches: 62

Results on Baseline - ResNet 50




```
Epoch [15/15]
  Train Loss: 0.9072, Train A
  Val Loss: 0.5325, Val Accur
```



DeIT Model by Facebook Research

***Model Used:** DeiT (Data-Efficient Image Transformer) Base Patch16-224*

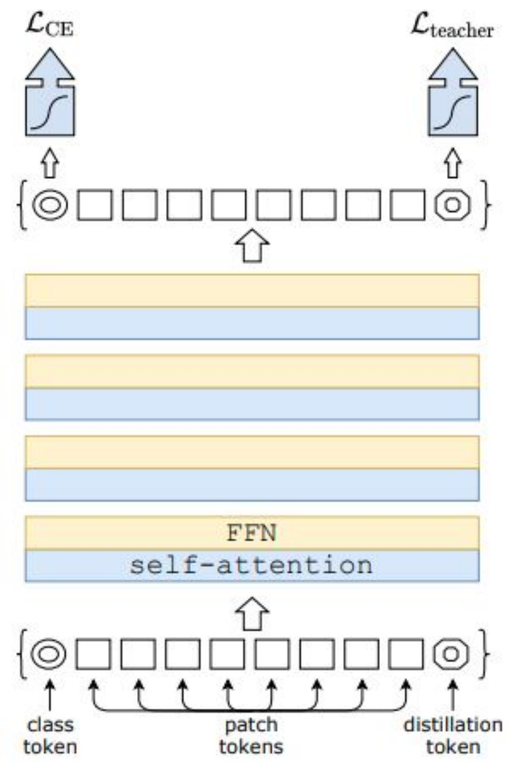
- **Patch Embedding:** Converts input images into patch tokens.
- **Self-Attention Mechanism:** Captures global relationships in image patches.
- **Head Adjustment:** Final layer replaced to match **[number of classes]**.

Advanced Augmentations:

- **Gaussian Blur:** Reduces noise while augmenting diversity.
- **CutMix:** Combines image pairs with interpolated labels for regularization.
- **Random Erasing:** Randomly masks parts of an image to enhance robustness.

Training Strategy:

- **Optimizer:** AdamW with learning rate **3e-4**, weight decay **1e-4**.
- **Loss Function:** Focal Loss to handle class imbalances effectively.
- **Learning Rate Scheduler:** Cosine Annealing Warm Restarts for smooth transitions.



DeIT Model Fine Tuning

Fine-Tuning Strategy:

- Earlier layers like **patch_embed** and **blocks** unfrozen for task-specific adaptation.
- Increased dropout rate (**0.3**) to combat overfitting.

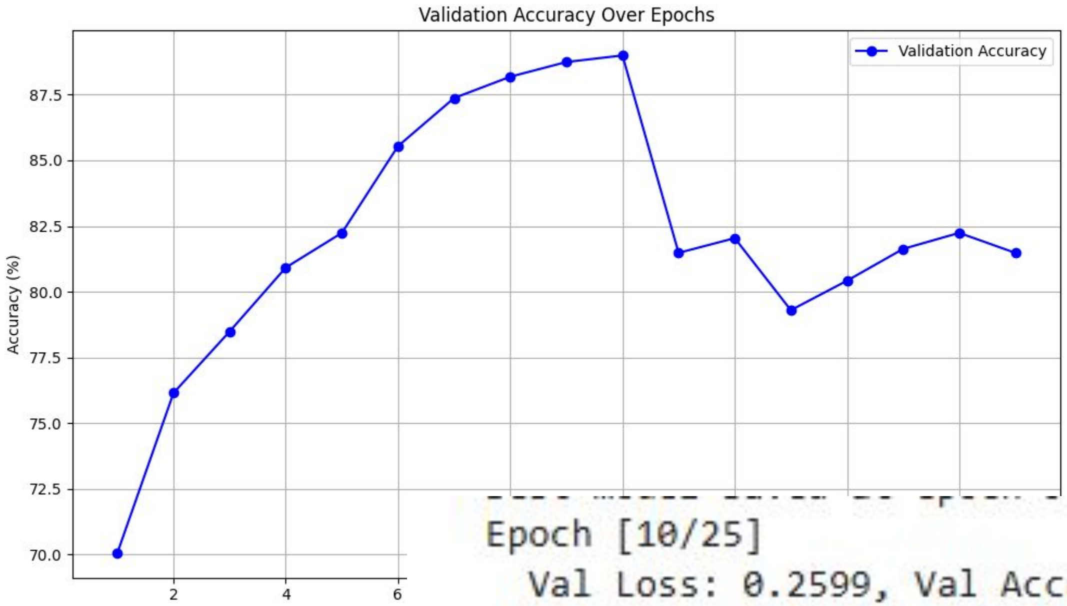
Mixed Precision Training:

- Utilized **GradScaler** for efficient GPU memory use and faster training.

Key Improvements:

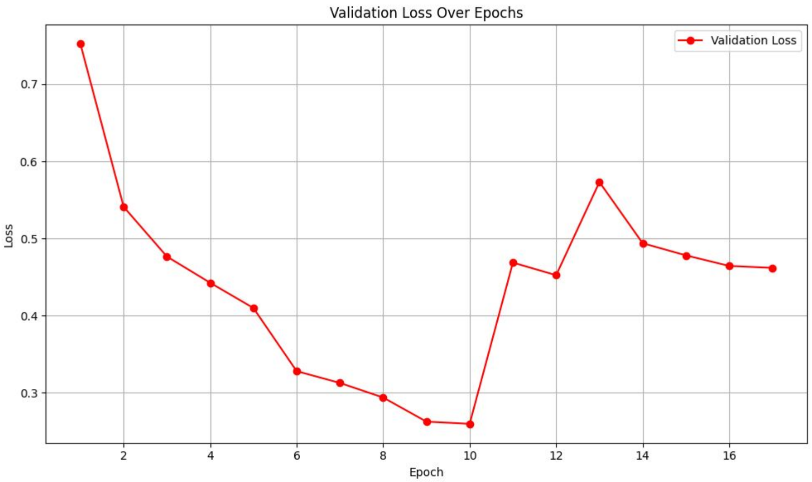
- Early stopping integrated to avoid overfitting.
- Achieved best validation F1-score using augmentations and CutMix.

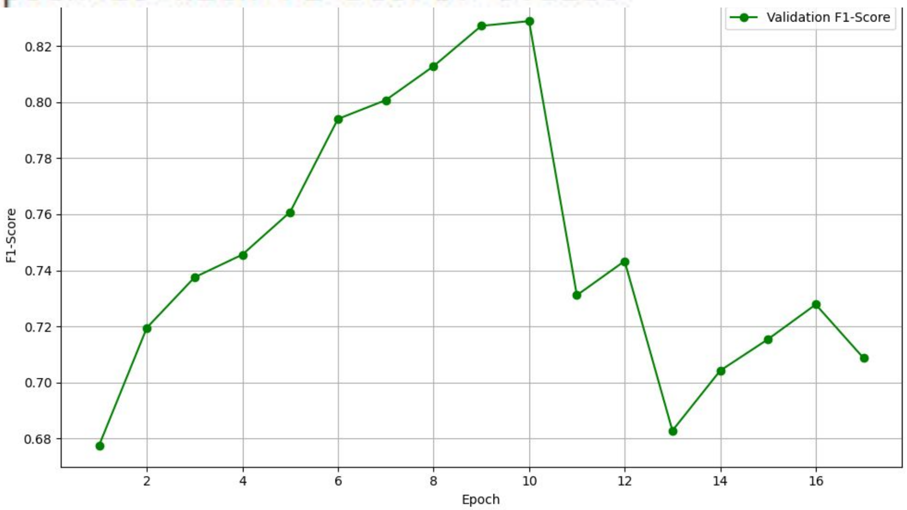
Results on DeIT Model



Model stopped training at epoch 17due to no improvement for 7 consecutive epochs

```
Epoch [10/25]
  Val Loss: 0.2599, Val Accuracy: 88.99%, Val F1: 0.8289
Best model saved at epoch 10 with F1 Score: 0.8289
```





ResNet 50 - Key Challenges and Solutions

Challenges:

- Handling corrupted or missing image files.
- Balancing learning rates for different layers.
- Preventing overfitting with limited labeled data.

Solutions:

- Filtered corrupted samples dynamically.
- Separate learning rates for fully connected and convolutional layers.
- Applied data augmentation for diversity.
- Layer Freezing:
  Balances general feature retention and task-specific learning.
  Unfreezing layer4 improved feature refinement and boosted accuracy.

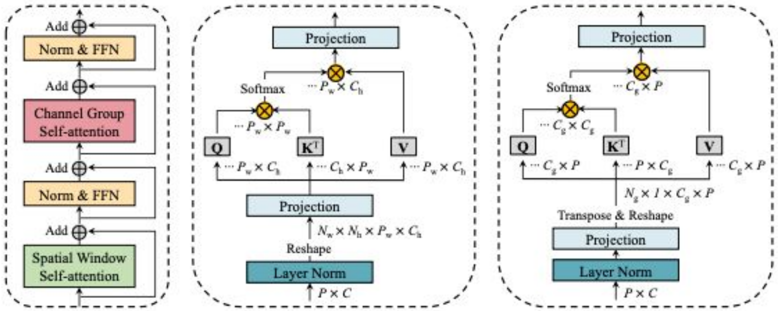
DaViT Model

Model Used: *DaViT (Dual Attention Vision Transformer)*

- davit_base

Why DaViT?

- Architecture is meant for image classification tasks,
  - Captures local and global context within image
- Best blend of CNNs and ViTs using dual attention
- Versatile for images with varying feature densities



(a) Dual Attention Block (b) Spatial Window Multihead Attention (c) Channel Group Attention

Augmentations:

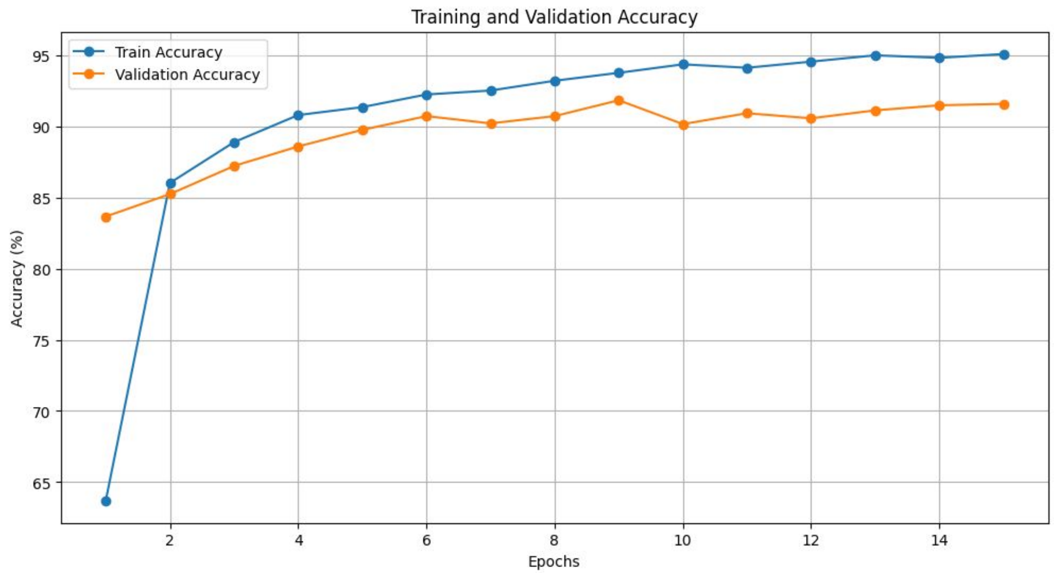
- **Random Augment:** Introduce variance into training data
- **Random Horizontal Flip:** Default given by base model

Training Strategy:

- **Optimizer:** AdamW with learning rate **1e-4**, for head parameters; learning rate **1e-5** and weight decay **1e-4** for backbone parameters
- **Loss Function:** Standard Cross Entropy Loss for classification tasks
- **Learning Rate Scheduler:** Cosine Annealing Warm Restarts for smooth transitions.

Results on DaViT Model


```
Epoch [15/15]
  Train Loss: 0.1825, Train Accuracy: 95.08%
  Val Loss: 0.3320, Val Accuracy: 91.58%
```



Future Work

- Ensemble Methods to Improve Accuracy
- Pseudo-labeling with Unlabeled Dataset
- Fine-tuning more layers or Hyperparameters
- Utilizing additional data augmentation strategies

Task Distribution

Nick Wang

Data Preprocessing

DaViT experiments

Slides

Model Research

Akashleena Sarkar

ResNet 50 fine tuning

DeIT model experiments

Slides

Model Research

Planning Work for Year 1

[Roundtable]

Concluding Remarks

[RICARDO]

[Remaining slides are some summary material that can be used, if needed]

Alianza MX Strategic Research Project

From Fields to Tables: Robotic Solutions for Sustainable Berry Harvesting

Ricardo Sanfelice
Electrical & Computer Engineering
UC Santa Cruz

Stefano Carpin
Electrical Engineering & Computer Science
UC Merced

Manuel Navarro-Gutiérrez
Mechatronics Engineering
Tecnológico de Monterrey

Jorge Isaac Chairez Oria
Mechatronics Engineering
Tecnológico de Monterrey

Collaboration:

Avoberrys el Valle

GRUPO LOS CERRITOS berries & aguacates

UC SANTA CRUZ UCMERCED Tecnológico de Monterrey

Project Overview

Goal: Motivated by the need to address agricultural labor shortages and inefficiencies, this project aims to develop robotic solutions for efficient and sustainable berry harvesting in greenhouses in Mexico and California. By utilizing robots equipped with sophisticated sensors and algorithms, this project intends to automate the berry picking process, ensuring precise identification and collection of ripe berries, thereby enhancing efficiency, reducing waste, and improving overall profitability for farmers.

Objectives:

1. Create databases of raspberries and blackberries growing in Mexican and American greenhouses.
2. Develop algorithms for identifying diseases in plants and berries.
3. Develop algorithms for detecting and locating berries through a computer vision system, as well as its maturity stage.
4. Integrate the computer vision system with a robotic arm for giving it target locations to reach.
5. Design hybrid control algorithms for driving the robotic arm to reach and collect berries in an efficient and safe manner.
6. Analyze resource allocation policies for efficient processing of harvested berries.
7. Reduce the time spent on bringing buckets of harvested berries to the quality control location.

Image credit: Washington Post

Image credit: strawberryplants.org

Image credit: Washington Post

Project Overview

Key Research Activities:

- Design image collection methodology
- Construct databases of raspberries and blackberries images in greenhouses
- Design artificial vision algorithms for berries detection
- Integration of artificial vision system information with robotic arms
- Design and test hybrid control algorithms on robotic arms
- Demonstration of technological prototypes in “Expo Agricola Jalisco”

Image credit: ScienceBusiness.net

Stakeholder Engagement and Impact

- Partners: Avoberrys el Valle, Grupo Los Cerritos
- Engagement: Collaborative research and technology demonstration with berry producers in California and Mexico.
- Impact: Enhanced efficiency in berry harvesting, reduced labor dependency, and promotion of sustainable practices.

Avoberrys el Valle

GRUPO LOS CERRITOS berries & aguacates ®

Next Steps and Future Plans

Next Steps

- Collaborator Engagement
- Prototype Development
- Demonstration and Outreach
- Pursue Additional Funding
- Training and Capacity Building
- Strategic Partnership Expansion

Image credit: business-humanrights.org

Scaling Up

- **Demonstration Farms**: Establishment of pilot farms in California and Mexico to showcase the developed robotic systems in real-world conditions.
- **Strategic Partnerships**: Expansion of collaborations with additional berry producers, agricultural associations, and technology companies to facilitate the widespread adoption of these robotic solutions.
- **Government Grants and Programs:** Pursue funding opportunities from agencies such as the USDA, NSF, and CONACYT to support further research, prototype development, and technology refinement.