ROBOTICS SOLUTIONS TO STRAWBERRY HARVESTING (AND LOGISTICS)

Prof Marc Hanheide

(with materials from Grzegorz Cielniak, Raymond Kirk and Sariah Mghames and others)

Lincoln Centre for Autonomous Systems (L-CAS)
University of Lincoln, UK
Challenges in Soft Fruit Production

Reliance of manual labour due to complex operations

Shrinking workforce due to socio-political pressures and ageing population
Thorvald - a modular mobile platform
Dedicated Strawberry Production Site

Realistic environment for data collection, algorithm development and field testing
RASberry: Robotic and Autonomous Systems for Berry Production

- Field Logistics and Autonomy (also working with Humans)
- Fruit Perception (for picking, phenotypic and yield prediction)
- Autonomous Picking Solutions
Field Logistics and Autonomy
Navigation and Autonomy

hybrid GPS & laser-based navigation

Realistic simulation environments

Automated tuning of parameters

(ICRA 2020)
Fleet Coordination

Avoiding deadlocks and moving robots in the field efficiently.

- Discrete Event Simulator (DES)
- Picker and robot agents with discrete states and transitions
- Navigation environment is discretised
- Really fast simulations compared to continuous-time simulations

RASberry DES Visualisation
Normal Operation

- A picker uses “Call-A-Robot” to request for a robot
- The coordinator adds the call to a task queue
- During processing of the task,
  - The coordinator assigns an idle robot to the picker
  - The robot goes to the picker and collect full trays
  - The picker can set “tray loaded” using “Call-A-Robot”
  - The robot goes to storage and trays are unloaded
  - The robot returns to its base station
Fleet Coordination - DES

- Anticipatory scheduling (predicting when the robot might be needed)
- Analysis of scalability in real-world environments (#pickers vs. #robots)
- Useful for global trajectory planning for individual robots
Perception of Humans

Motivations:

• Safety during navigation
• Approach or avoid workers in the field

Challenges:

• Sensor noise, dirt
• Occlusions, lighting, misclassification
• Wide open areas, uneven terrain
Sensors

- Stereo camera
- Thermal camera
- 3D LiDAR
- RGB-D camera
Human Detection and Tracking

- Independent sw modules for **human detection** are assigned to each sensor.
- Detections are transformed from (local) sensor to (global) robot frame of ref.
- Finally combined in a multisensor/multiperson Bayesian estimator (i.e. UKF)
Examples in the Field

Single person

Multiple people
Fruit Perception and Tracking
Robotic Perception for the Soft Fruit Industry

Research Question

- Where are the berries?
- Where are the berries in real world coordinates?
- How can we harvest the berries?
- Can we harvest berries we can’t see?
- Have I seen this berry before?

Solution

- Detecting Berries From Images. Enables Analysis.
- Enables Harvesting Applications and Virtual Agronomy.
- Reduces labor demand, damage to berries, increases harvest throughput.
- Improves detection and harvesting rates.
- Fruit count information, maturity maps, and yield estimation.

Industry/Grower Impact
Technical Challenges

Outdoors - changing lighting/weather conditions

Natural variation of fruit/plants in appearance, shape, orientation

Temporal changes and metamorphosis due to growth

Occlusions by other berries, leaves, trellis

Limitations of the current sensing technology
Problem

Environmental features outdoors (sunlight intensity, weather conditions, etc.) affect perceptual appearance of objects and consequently the performance of detection systems.

Solution

We fuse RGB data with bio-inspired features to classify produce over a greater number of conditions.
Human-Inspired Features

Colour Opponency
Data Collection (summer 2019)

<table>
<thead>
<tr>
<th>Viewpoint</th>
<th>Training</th>
<th>Validation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>120 (80%)</td>
<td>10 (6.6%)</td>
<td>130</td>
</tr>
<tr>
<td>V2</td>
<td>0 (0%)</td>
<td>10 (6.6%)</td>
<td>10</td>
</tr>
<tr>
<td>V3</td>
<td>0 (0%)</td>
<td>10 (6.6%)</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>120 (80%)</td>
<td>30 (20%)</td>
<td>150</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bounding Boxes</th>
<th>Ripe</th>
<th>Unripe</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>673</td>
<td>2680</td>
<td>3353</td>
</tr>
<tr>
<td>Validation</td>
<td>217</td>
<td>649</td>
<td>886</td>
</tr>
<tr>
<td>Total</td>
<td>890</td>
<td>3329</td>
<td>4219</td>
</tr>
</tbody>
</table>
Fruit Detection
Results

<table>
<thead>
<tr>
<th>Class</th>
<th>View</th>
<th>Score</th>
<th>RGB</th>
<th>CIE Lab</th>
<th>Early Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both Classes</td>
<td>V₁</td>
<td>F₁</td>
<td>0.744</td>
<td>0.710</td>
<td>0.747</td>
</tr>
<tr>
<td>Both Classes</td>
<td>V₁</td>
<td>AP</td>
<td>0.722</td>
<td>0.695</td>
<td>0.748</td>
</tr>
<tr>
<td>Both Classes</td>
<td>V₁</td>
<td>AR</td>
<td>0.879</td>
<td>0.844</td>
<td>0.909</td>
</tr>
<tr>
<td>Both Classes</td>
<td>V₂₋₃</td>
<td>F₁</td>
<td>0.682</td>
<td>0.622</td>
<td>0.704</td>
</tr>
<tr>
<td>Both Classes</td>
<td>V₂₋₃</td>
<td>AP</td>
<td>0.659</td>
<td>0.586</td>
<td>0.694</td>
</tr>
<tr>
<td>Both Classes</td>
<td>V₂₋₃</td>
<td>AR</td>
<td>0.812</td>
<td>0.761</td>
<td>0.851</td>
</tr>
<tr>
<td>Ripe Strawberry</td>
<td>V₁</td>
<td>F₁</td>
<td>0.883</td>
<td>0.825</td>
<td>0.897</td>
</tr>
<tr>
<td>Ripe Strawberry</td>
<td>V₁</td>
<td>AP</td>
<td>0.816</td>
<td>0.571</td>
<td>0.678</td>
</tr>
<tr>
<td>Ripe Strawberry</td>
<td>V₁</td>
<td>AR</td>
<td>0.807</td>
<td>0.767</td>
<td>0.892</td>
</tr>
<tr>
<td>Ripe Strawberry</td>
<td>V₂₋₃</td>
<td>F₁</td>
<td>0.697</td>
<td>0.662</td>
<td>0.729</td>
</tr>
<tr>
<td>Ripe Strawberry</td>
<td>V₂₋₃</td>
<td>AP</td>
<td>0.658</td>
<td>0.621</td>
<td>0.719</td>
</tr>
<tr>
<td>Ripe Strawberry</td>
<td>V₂₋₃</td>
<td>AR</td>
<td>0.806</td>
<td>0.777</td>
<td>0.877</td>
</tr>
<tr>
<td>Unripe Strawberry</td>
<td>V₁</td>
<td>F₁</td>
<td>0.805</td>
<td>0.795</td>
<td>0.797</td>
</tr>
<tr>
<td>Unripe Strawberry</td>
<td>V₁</td>
<td>AP</td>
<td>0.828</td>
<td>0.819</td>
<td>0.818</td>
</tr>
<tr>
<td>Unripe Strawberry</td>
<td>V₁</td>
<td>AR</td>
<td>0.933</td>
<td>0.922</td>
<td>0.927</td>
</tr>
<tr>
<td>Unripe Strawberry</td>
<td>V₂₋₃</td>
<td>F₁</td>
<td>0.663</td>
<td>0.582</td>
<td>0.679</td>
</tr>
<tr>
<td>Unripe Strawberry</td>
<td>V₂₋₃</td>
<td>AP</td>
<td>0.658</td>
<td>0.552</td>
<td>0.668</td>
</tr>
<tr>
<td>Unripe Strawberry</td>
<td>V₂₋₃</td>
<td>AR</td>
<td>0.819</td>
<td>0.745</td>
<td>0.825</td>
</tr>
</tbody>
</table>

Table 7. Performance of the Early Fusion Network on a Nvidia GTX 1080 Ti 11GB (single forward pass).

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Model Inference Time</th>
<th>Frames Per Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920 x 1080</td>
<td>0.073 s</td>
<td>15.71</td>
</tr>
<tr>
<td>1280 x 720</td>
<td>0.038 s</td>
<td>26.33</td>
</tr>
</tbody>
</table>

(a) RGB network detection showing failure cases.

(b) Early Fusion network detection showing improved results.
Not just for strawberries!
### SOTA in DL for Fruit Detection

<table>
<thead>
<tr>
<th>Method</th>
<th># Images</th>
<th>Availability</th>
<th>Viewpoint</th>
<th>Multi Spectra</th>
<th>Controlled</th>
<th>Natural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yu et al. [34]</td>
<td>1900</td>
<td>×</td>
<td>Side on (Close)</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Chen et al. [35]</td>
<td>12526</td>
<td>×</td>
<td>Aerial</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Lamb and Chuah [36]</td>
<td>4550</td>
<td>×</td>
<td>Ground</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Ge et al. [37]</td>
<td>-</td>
<td>×</td>
<td>Side on</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Sa et al. [8]</td>
<td>122</td>
<td>✓</td>
<td>Side on</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>L* a<em>b</em> Fruits (Ours)</td>
<td>150</td>
<td>✓</td>
<td>Multiple</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Network</th>
<th>AP (IoU 0.5)</th>
<th>F₁ (IoU 0.5)</th>
<th>Inference Speed (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yu et al. [34]</td>
<td>Mask R-CNN - ResNet-50</td>
<td>-</td>
<td>-</td>
<td>0.13 @ 640 × 480 px</td>
</tr>
<tr>
<td>Chen et al. [35]</td>
<td>Faster R-CNN - ResNet50</td>
<td>0.77</td>
<td>-</td>
<td>0.11 @ 480 × 380 px</td>
</tr>
<tr>
<td>Lamb and Chuah [36]</td>
<td>Single Shot Detector (SSD)</td>
<td>0.84</td>
<td>-</td>
<td>0.61 @ 360 × 640 px</td>
</tr>
<tr>
<td>Ge et al. [37]</td>
<td>Mask R-CNN - ResNet-101</td>
<td>0.81</td>
<td>0.90</td>
<td>0.62 @ 640 × 480 px</td>
</tr>
<tr>
<td>Sa et al. [8]</td>
<td>Faster RCNN - VGG-16</td>
<td>-</td>
<td>0.79</td>
<td>0.39 @ 1296 × 964 px</td>
</tr>
<tr>
<td>L* a<em>b</em> Fruits (Ours)</td>
<td>RetinaNet, ResNet-18</td>
<td>0.75</td>
<td>0.75</td>
<td>0.07 @ 1920 × 1080 px</td>
</tr>
</tbody>
</table>
Yield Prediction (Fruit Counting)
Picking
Fruit Localisation (3D projection)

3D Projection of Strawberry Detections Showing Accurate Localisation
Harvesting Localisation Failures

Pros:
Current detection system can now successfully pick single berries with good separation.
Fast 3D detections.

Cons:
Cannot pick berries in clusters due to unforeseen obstructions.
No understanding of the harvesting apparatus perspective.
Picking (Path Planning for Clusters)

- Probabilistic Motion Primitives to deal with Clusters
- submitted to IROS 2020
System Integration

- ROS-based Software Infrastructure

RASberry packages

L-CAS robotic software stack
Summary and Future Work

• RASberry - fleet of autonomous robots for soft fruit industry
• Various applications: in-field logistics, plant care, yield prediction, fruit picking
• An integrated challenge:
  • logistics needed after picking
  • perception for picking and yield prediction, etc.
  • safely deal with humans in the field