ML for perception and navigation in robotics

Alessandro Giusti
Dalle Molle Institute for Artificial Intelligence
Lugano, Switzerland

USI - SUPSI
Outline

1. Vision-based roughness estimation of metal surfaces
2. Vision-based proximity Human-Drone Interaction
3. Augmented reality for microscopy applications
1. VISION-BASED ROUGHNESS ESTIMATION OF METAL SURFACES
Visual measurement of surface roughness

Image

Image Processing Software

Results

Ra = 0.98 ± 0.13 µm
## Machine Learning Approach (1)

### Training set

<table>
<thead>
<tr>
<th>Image</th>
<th>Known Ra</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>0.17</td>
</tr>
</tbody>
</table>
## Machine Learning Approach (2)

### Training set

<table>
<thead>
<tr>
<th>Image</th>
<th>Known Ra</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image 1" /></td>
<td>2.09</td>
</tr>
<tr>
<td><img src="image2.png" alt="Image 2" /></td>
<td>1.21</td>
</tr>
<tr>
<td><img src="image3.png" alt="Image 3" /></td>
<td>0.64</td>
</tr>
<tr>
<td><img src="image4.png" alt="Image 4" /></td>
<td>0.37</td>
</tr>
<tr>
<td><img src="image5.png" alt="Image 5" /></td>
<td>0.17</td>
</tr>
</tbody>
</table>

![Diagram](diagram.png)
## Machine Learning Approach (3)

**Training set**

<table>
<thead>
<tr>
<th>Image</th>
<th>Known Ra</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>0.17</td>
</tr>
</tbody>
</table>

![Regressor](image)
Machine Learning Approach (4)

Training set

<table>
<thead>
<tr>
<th>Image</th>
<th>Known Ra</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>0.17</td>
</tr>
</tbody>
</table>

New image input

Regressor
Machine Learning Approach (5)

Training set

<table>
<thead>
<tr>
<th>Image</th>
<th>Known Ra</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>0.17</td>
</tr>
</tbody>
</table>

New image input

Regressor

Ra = 0.69 ± 0.11 µm

Output
Deep Neural Networks as Regressors

**Input layer**
- a grid of 64x64 neurons
- (raw image pixel intensities)

**Convolutional and Max-pooling layers**
- a standard lenet-like architecture for feature extraction

**Two fully-connected layers**
- as a regressor

**One output neuron**
- activation representing the Ra value directly
Example results
Example results
Example results
Cross validated results on cavities
Anomaly Detection Approaches

- Large defect-free datasets for various Ra
- No annotated data available for defects
- Wide variability of possible anomalies:
  - imaging issues (focus/illumination)
  - surface defects
Example input
Anomaly map
Estimated Ra map
Estimated Ra map in non-anomaly regions

Ra valid 0.28
Output, user-facing interface
Detection of localized defects
Detection of localized defects
Detection of localized defects
2. VISION-BASED PROXIMITY HUMAN-ROBOT INTERACTION

Videos, links, publications and code:
https://github.com/idsia-robotics/proximity-quadrotor-learning
Control a drone flying close to humans
Easy! (using motion tracking)
Proximity control using motion tracking
Goal: vision-based control
A deep net for end-to-end visual control

Camera input

Pitch
Roll
Yaw vel
Z vel
Datasets
Prediction performance

• ground truth control
• predicted values (by three different models)
Comparison with ground truth control

ground truth control

vision-based control

Original controller

A1

We repeat the experiment multiple times, and compare the resulting trajectories
Control performance and robustness
Domain randomization for generalization
Preliminary Results

- Trained without domain randomization
- Trained with domain randomization
3. ARTIFICIAL DEFOCUS FOR DISPLAYING MARKERS IN MICROSCOPY Z-STACKS

Flea chest – original image stack courtesy Daniel Stoupin – Used with permission
Human embryo at 4-cell stage – validation and correction of automated segmentation
Summary

The talk covered three topics

1. Vision-based roughness estimation of metal surface
2. Vision-based proximity Human Robot Interaction
3. Augmented reality for microscopy applications

Contact: alessandrog@idsia.ch
https://idsia-robotics.github.io/