Estimation of relative position between objects for robotic insertion of LDOs using classification learning methods

TECHNION Israel Institute of Technology

Sher Hazan
Anath Fischer, Ronit Schneor

CAD Laboratory
Mechanical Engineering Faculty

Abstract

This work proposes a robust method for evaluating the relative position between LDOs in real-time assembly tasks. The proposed approach utilizes a classification CNN architecture with RGB-D data fusion. Several fusing approaches were tested and compared. In order to use classification CNN on a continuous problem, the scene space was divided into sub-areas as classes. The classification resulted with the prediction of probabilities of each class. Subsequently, these probabilities were combined using linear interpolation, resulting in an accurate evaluation of the relative position between the objects. For the validation process, three Peg-In-Hole tasks were defined. An RGB-D realistic database was created for each task. Then, pre-processing and augmentation techniques were applied on the data. Finally, the CNNs were trained separately using the same setup to ensure comparable results. The method was evaluated by embedding and testing it in the insertion process of each task at the Technion and the in industry. The feasibility of the proposed method was demonstrated in those tasks, showing an automatic process with high performance.

Problem Statement

Estimating the relative position between objects in PIH tasks.
With a focus on linear deformable objects (LDO).
Using classification learning methods.

Challenges:

Real time operation: 30 [FPS].

Accuracy:

- 1 [mm] in the radius
- 10 [degrees] in the angle.

Robustness: To environment changes and camera viewpoints.

Approach

A pipeline for estimating the relative position between objects using learning method embedded into real robotic cell:

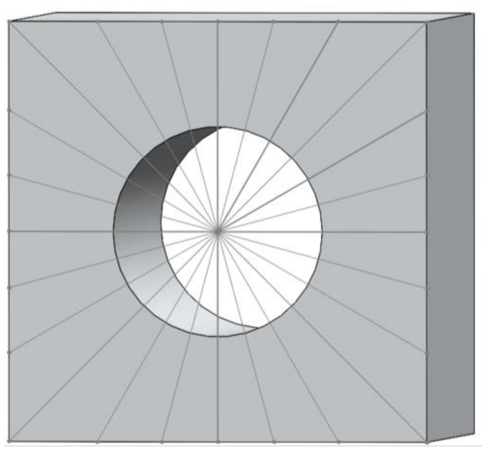
Capturing RGB-D data:

The RGB-D data captured on the robotic cell using RGB-D camera.

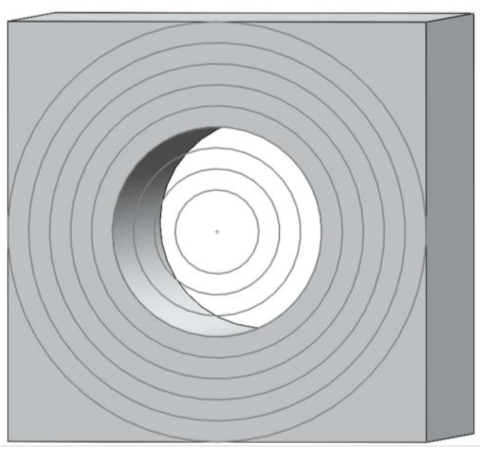
Estimation of relative position:

- **Discretization**: In order to use classification methods on a continuous tasks, the scene space was divided into sub-areas as classes.
- **Classification**: Classifying the interaction state between the hole and the object, according to the classes defined. resulting in the prediction of probabilities to be in each class.
- **Estimation**: The probability vector is combined using linear interpolation, resulting in an accurate evaluation of the relative position between the objects, in terms of radius and angle.

Dividing by angle



Dividing by radius

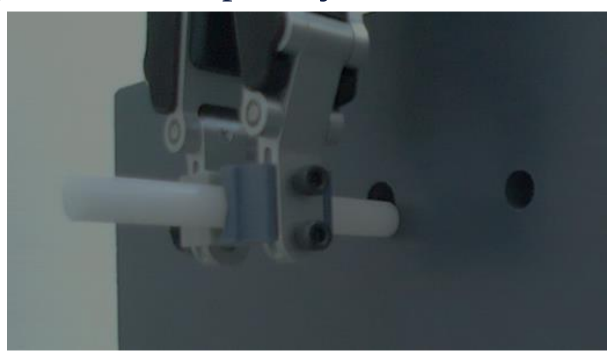


Test cases

For the validation process of the proposed method, three PIH tasks were defined, varying in their complexity.

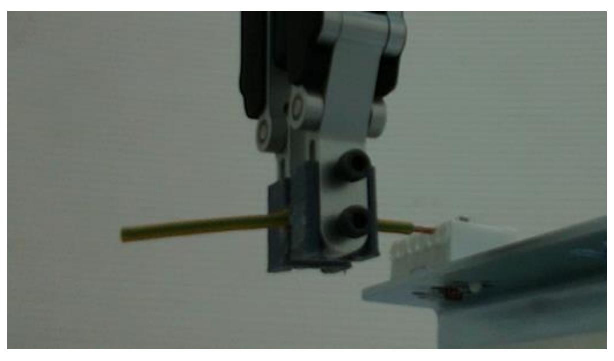
Classical Peg in Hole

Rigid – Rigid



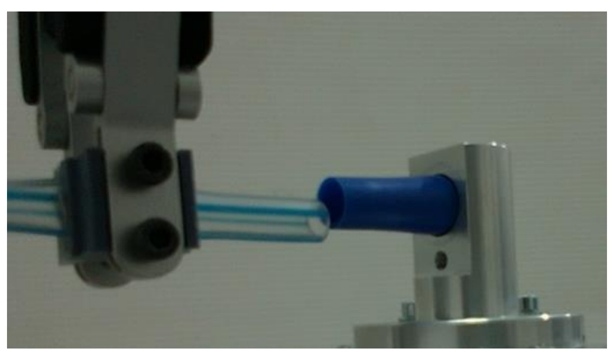
Wiring 1[mm] wire to a connector

Non-Rigid – Rigid

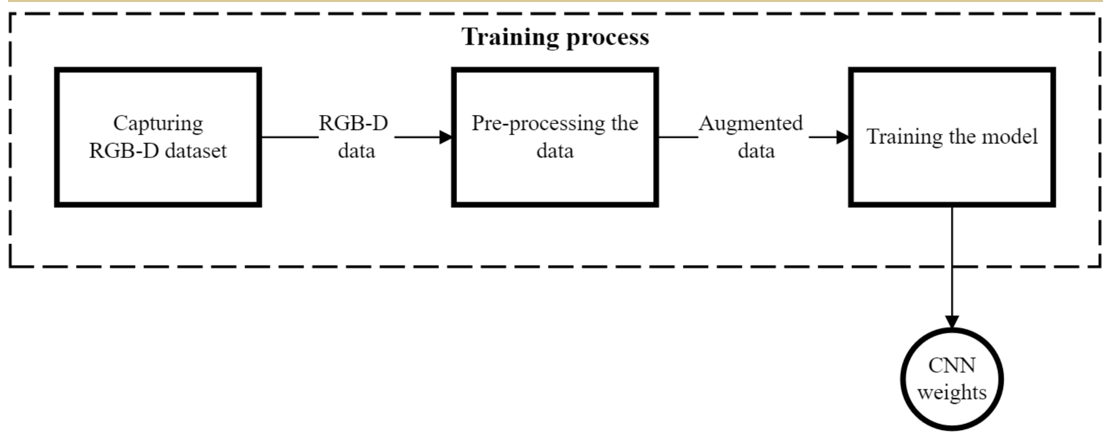


Insert medical pipe to non-rigid connector

Non-Rigid – Non-Rigid



Training process

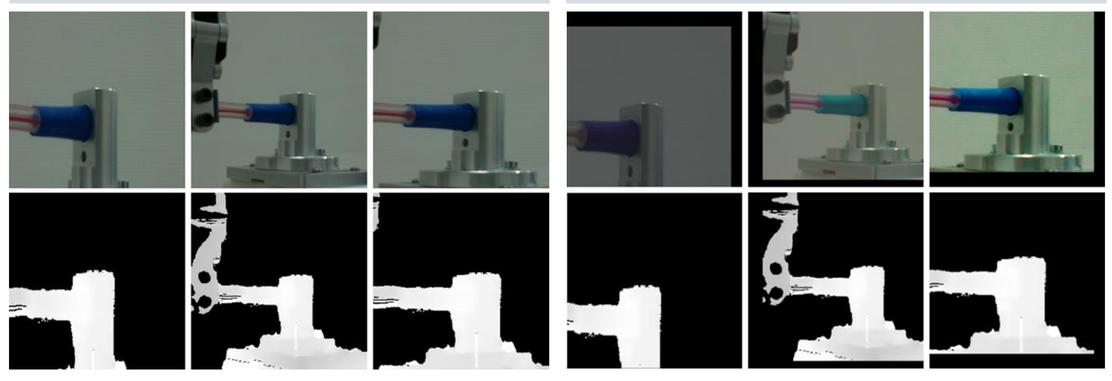


Capturing RGB-D dataset: 2500 RGB-D scenes was captured for each task.

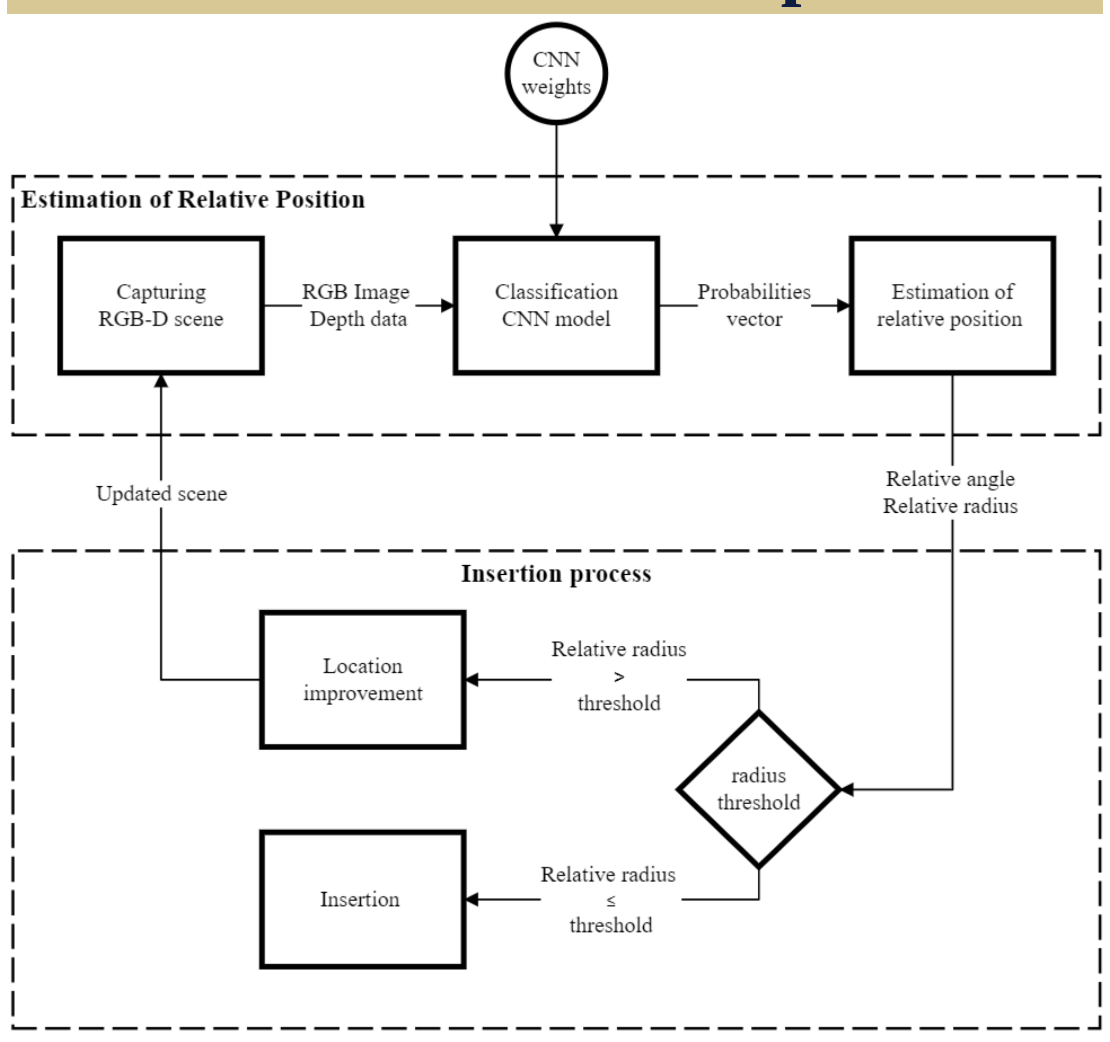
Pre-processing the data: Extract patches of various sizes and shapes from the dataset, to expand it to 7,500 unique scenes. Apply augmentations include lighting and color adjustments as well as translations and rotations.

Training the model: To ensure comparable results, all networks were trained using identical setup and hyperparameters.

Original images | Augmented images



Estimation of relative position



Capturing RGB-D scene: Capturing the RGB-D scene.

Rotated the scene 180 degrees. Resizing and cropping a 240x240 patch to match the required input.

Classification CNN models: For each task, five CNN architectures were designed and trained:

model	Input	Fusion location
RGB	RGB image	-
Depth	Depth data	-
RGB-D early fusion	Concatenated RGB-D	Before input
RGB-D late fusion	Separated RGB-D	Inside the model
RGB-D parallel fusion	Separated RGB-D	After output

Estimation of relative position: The continuous relative position α , r is obtained from the probability vector in three steps:

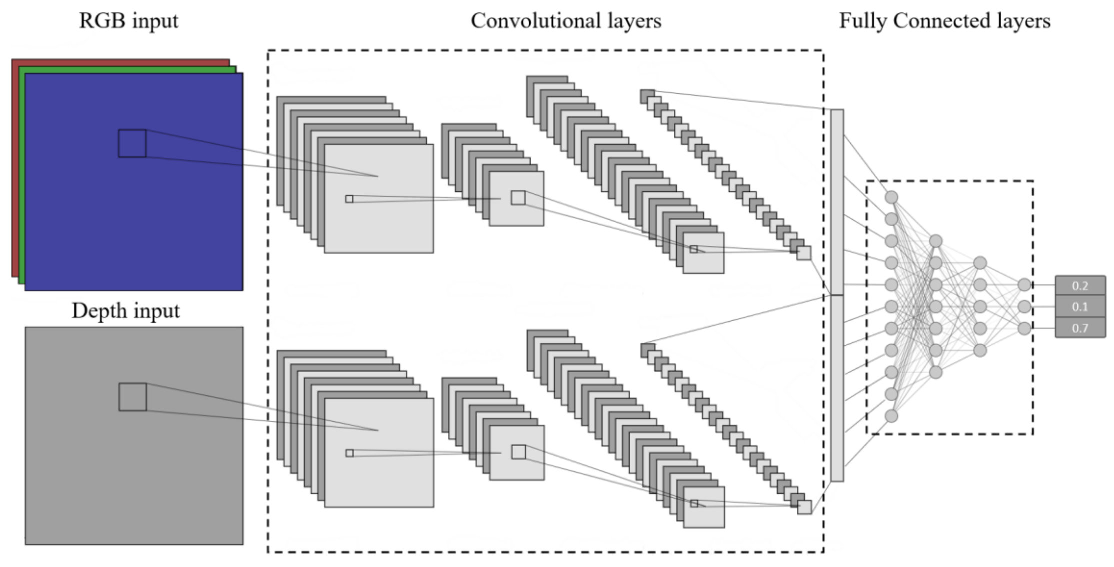
- Thresholding negligible values to suppress noise from the probability vector.
- Normalizing the filtered probability vector.
- Calculation of the mean value across all classes based on the probability vector.

Location improvement: If the relative radius is greater than the threshold The robotic arm improve its location.

Insertion: If the relative radius is smaller then the threshold The robotic arm start the insertion.

The threshold is determined based on the minimum error in the radius, which ensures successful insertion of the object.

Late fusion architecture

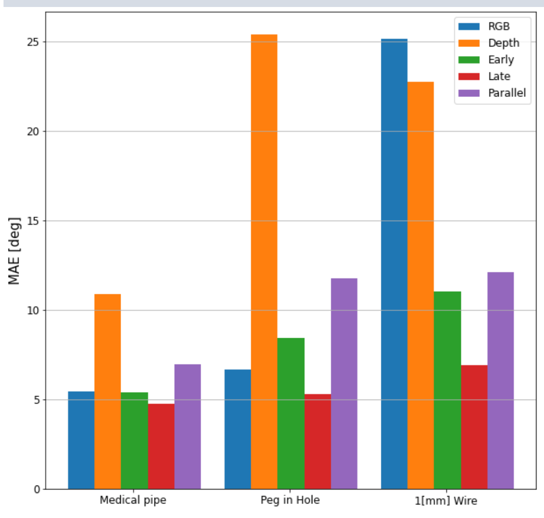


Performance Analysis

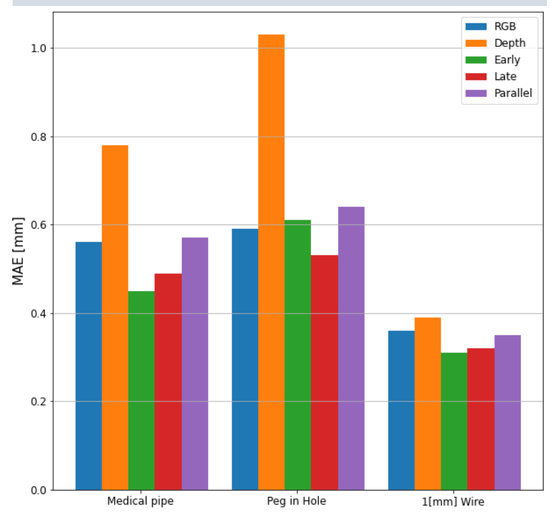
The method's effectiveness was evaluated based on the Mean Absolute Error (MAE) across the different models and tasks.

The late fusion method stands out as the most effective overall, consistently producing the lowest errors across most cases.

Angular MAE



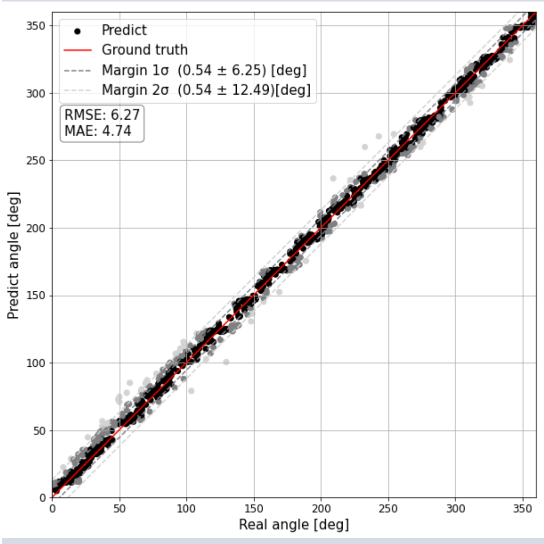
Radial MAR



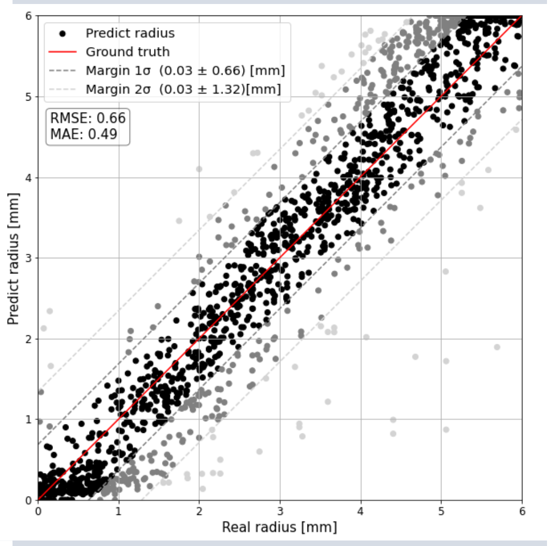
The following graphs represent the results of the late fusion model applied to the medical pipe task.

Medical pipe insertion with late fusion model

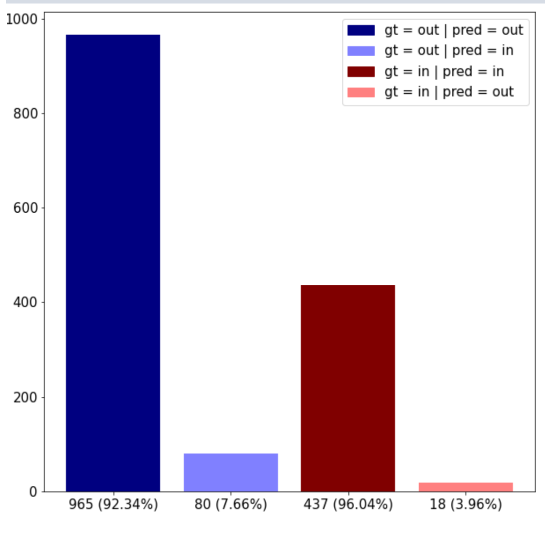
Real vs Predicted Radius



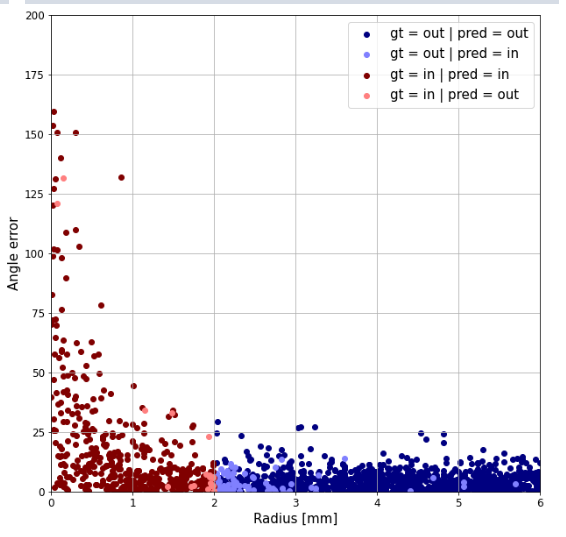
Real vs Predicted Angle



In Out prediction



Angle error vs Radius

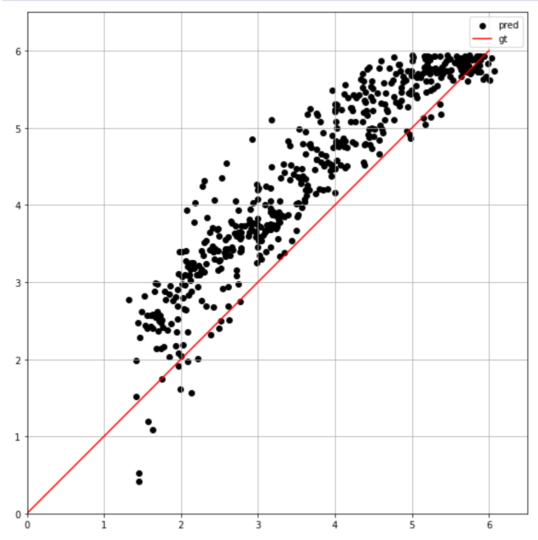


Industrial Application

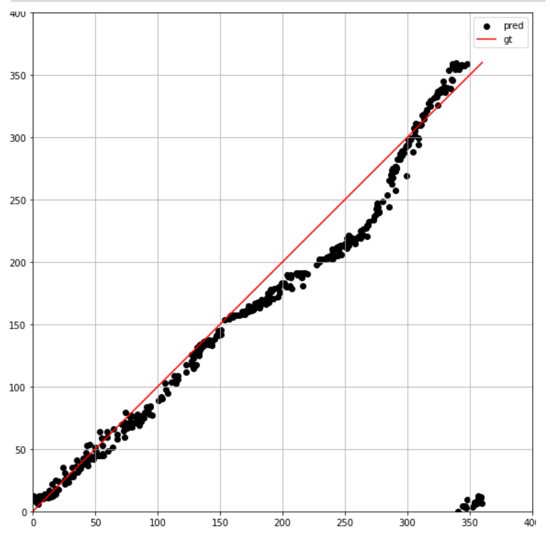
To validate the real-world applicability and the generalization of the method, a performance test was conducted in the industry. This test involved the task of medical pipe insertion and was carried out without additional training.

Medical pipe insertion at the industry

Real vs Predicted Radius



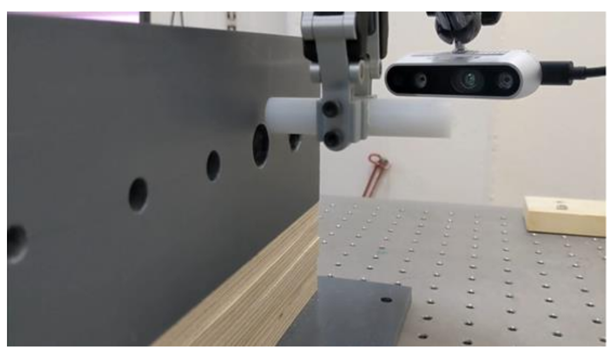
Real vs Predicted Angle



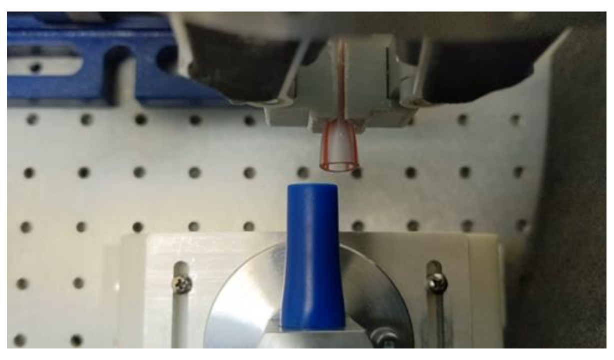
Video Demonstrations

Observe the insertion process in action by scanning the QR code:

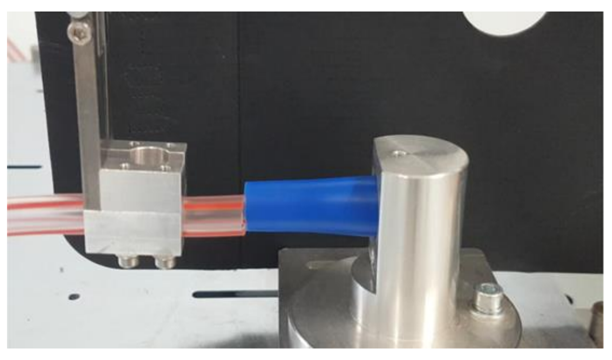
Peg in Hole insertion



Medical Pipe insertion



Medical Pipe Insertion at industry



Contact: Sher Hazan sherhazan1115@gmail

References:

[1] Dhruv, P. and Naskar, S., 2020. Image classification using convolutional neural network (CNN) and recurrent neural network (RNN): a review. Machine Learning and Information Processing: Proceedings of ICMLIP 2019, pp.367-381.

[2] Jiang, J., Huang, Z., Bi, Z., Ma, X. and Yu, G., 2020. State-of-the-Art control strategies for robotic PiH assembly. Robotics and Computer-Integrated Manufacturing, 65, p.101894.

[3] Gao, M.; Jiang, J.; Zou, G.; John, V.; Liu, Z. RGB-D-based object recognition using multimodal convolutional neural networks: A survey. IEEE Access 2019, 7, 43110–43136.

[4] Chen, J., Zhang, L., Liu, Y. and Xu, C., 2020, July. Survey on 6D pose estimation of rigid object. In 2020 39th Chinese Control Conference (CCC) (pp. 7440-7445). IEEE.