Poster Title: Robot Grasp Position Prediction Using Convolutional Neural Networks
Author: Nikhil Sawane, Yashwant Gandham
Course: CSCI 5922: Neural Nets and Deep Learning, Spring 2025

Introduction:

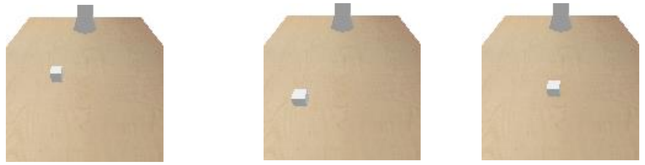
- **Objective:** Predict 3D grasp positions directly from RGB images using deep learning.
- **Motivation:** Classical planning struggles with uncertain, dynamic environments.
- **Approach:** Lightweight CNN model predicts (x, y, z) placement directly from images.

Problem Setup:

- **Simulation Environment:** PyBullet with Franka Panda robot arm.
- **Simplification:** Static conveyor and stationary objects.
- **Goal:** Predict (x, y, z) grasp position from input image.

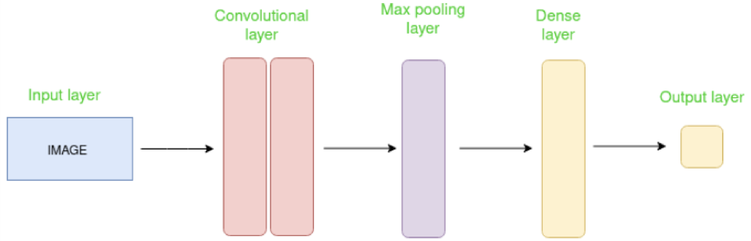
Dataset:

- **Images:** 100 simulated scenes (grasp_001.png to grasp_100.png).
- **Labels:** Randomized dummy labels (x, y, z, score).
- **Input Format:** 128x128 RGB images.
- **Output:** Predicted grasp position (x, y, z).



Methodology:

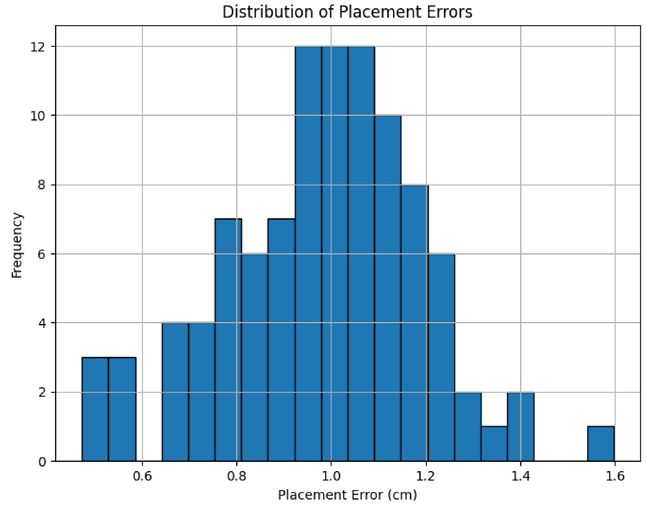
- **CNN Architecture:**
  - 3 Convolutional Layers (ReLU + MaxPooling)
  - Fully Connected Layers
- **Loss Function:** Mean Squared Error (MSE)
- **Training Status:** Evaluation with randomly initialized model (no real training yet).



Simple CNN architecture

Results:

- **Mean Placement Error:** 1.00 cm
- **Success Rate:** 100% (error threshold = 5 cm)



Conclusion and Future Work:

- Successfully implemented the full prediction pipeline from simulation to evaluation.
- Future steps:
  - Collect ground-truth grasp labels.
  - Deploy trained model in dynamic real-world setups.

Acknowledgments:

- CSCI 5922: Neural Nets and Deep Learning
- Thanks to course instructors and TAs for their support.