HPC Lunch and Learn

March 2019

Deep Learning Applications in Manned Spaceflight

Matthew Noyes

NASA JSC - Advanced Operational Concepts Lab

Lawrence Hessburg
Graduate Student

Lui Wang
AR Lead

Agenda

- Deep Learning Overview/Definitions
- Intelligent Personal Coach
  - Use of open source datasets for space applications
- Safety Analysis with Deep Learning
  - Use of open source datasets for space applications
- 6DOF Object Pose Estimation with Virtual Training Dataset
  - Doing deep learning with hard-to-generate datasets
- Potential Areas of Research/Future Needs for Space Applications

Intro - What is “Deep Learning”?

- **Artificial Intelligence** – a simulation of intelligent systems
- **Machine Learning** – Self-modifying AI
- **Deep Learning** – Self modifying AI with multiple hidden layers
- Neural Networks/Deep Learning attempt to replicate the structure of the human brain on a rudimentary level

What is “Training”?

- Basic steps of training:
  - Weights are initially randomized
  - Feed training examples through network and compute Loss at output
  - Reinforce “good” weights through backpropagation
  - Evaluate performance on test examples
  - Repeat
- Must have accurate Loss function and large, diverse, *labeled* training set

Useful Definitions

Loss - How far is the prediction from the truth?

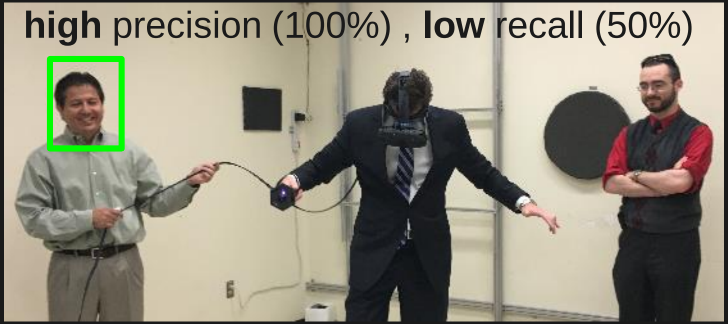
Precision - How many of the predictions are correct?

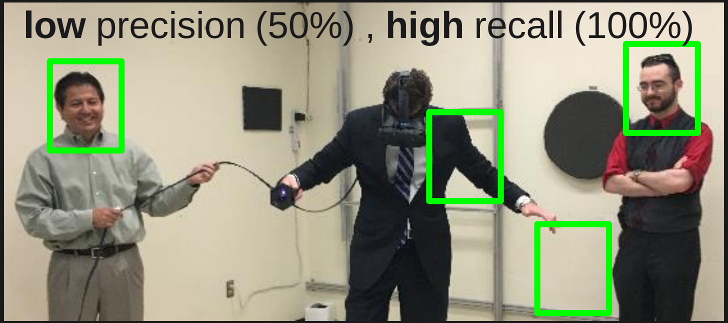
Recall - How many of the objects were found?

Mean Average Precision - Takes into account both Precision and Recall

Ground Truth - Known, labeled example

	apple	banana	lime	orange	lemon
truth	0	0	0	1	0
prediction	0.05	1e-4	0.08	0.7599	0.11





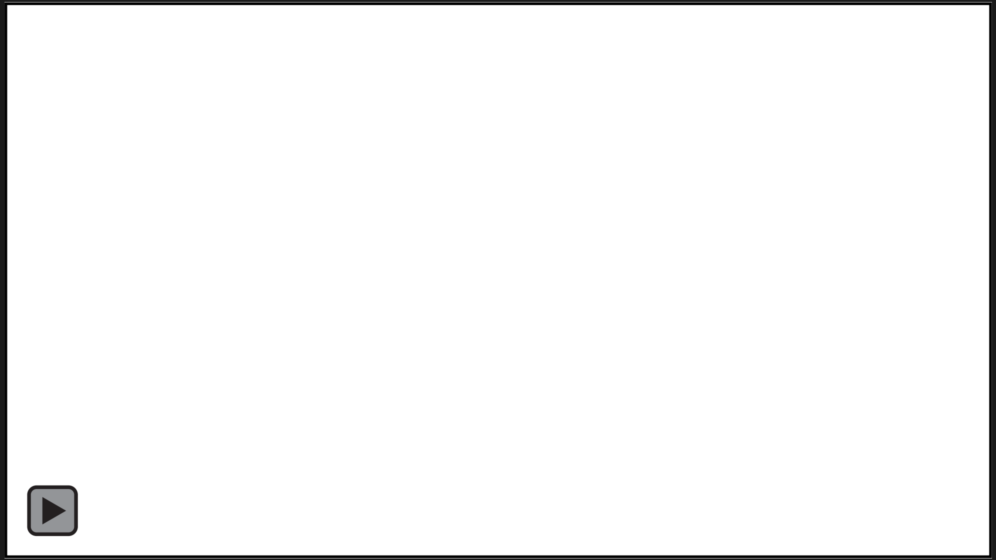
Intelligent Personal Coach

Intelligent Personal Coach - Background

- Exercise is vital
- Communication delay
- Characteristics of Intelligent Personal Coach:
  - Health/Performance monitoring
  - Workout/Lesson Planning
  - **Remediation** - real-time guidance during exercise

Intelligent Personal Coach - Requirements

- Real-time feedback
- Intuitive UI, easy to act upon
- **No wearables** - visual or audio feedback only
- Initially focused on simple squat exercise:
  - Keep upper back straight
  - Keep chest up
  - Keep correct bar path during range of motion



Intelligent Personal Coach - openpose

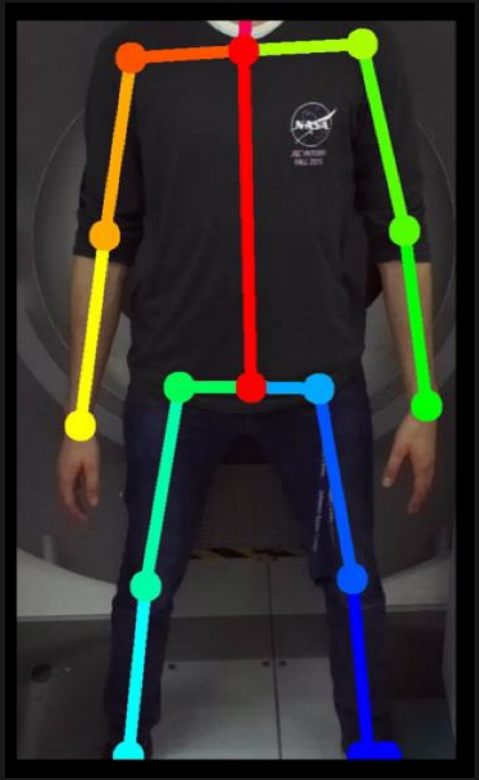
- Openpose: deep-learning based human pose estimation framework developed by CMU
- Accurate keypoint detection for desired joints
- Works on any image/video feed (i.e. webcam)
- But: joint keypoints in image-space isn’t enough to accurately judge exercise form

Intelligent Personal Coach - ZED camera

- Uses dual-camera parallax to estimate depth in an image/video
- Realtime depth map at full camera resolution
- Combine depth map with joint keypoints to determine form

Intelligent Personal Coach - BodyModel

- BodyModel class to integrate pixel-space keypoints from openpose with depth info from ZED camera.
- ZED video stream feeds to openpose network
- Underlying layer visualizations can be built upon

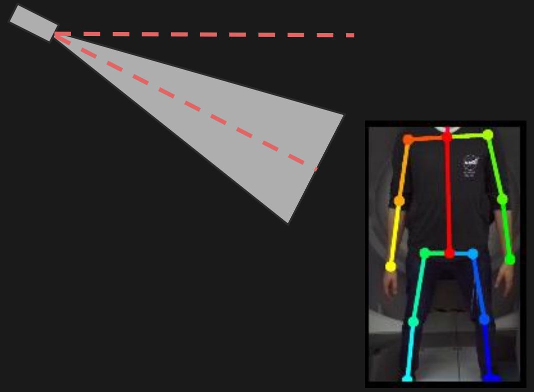


Intelligent Personal Coach - Design

- Single ZED camera mounted directly in front of subject on the ceiling.

- Short throw projector on opposing wall for visualization

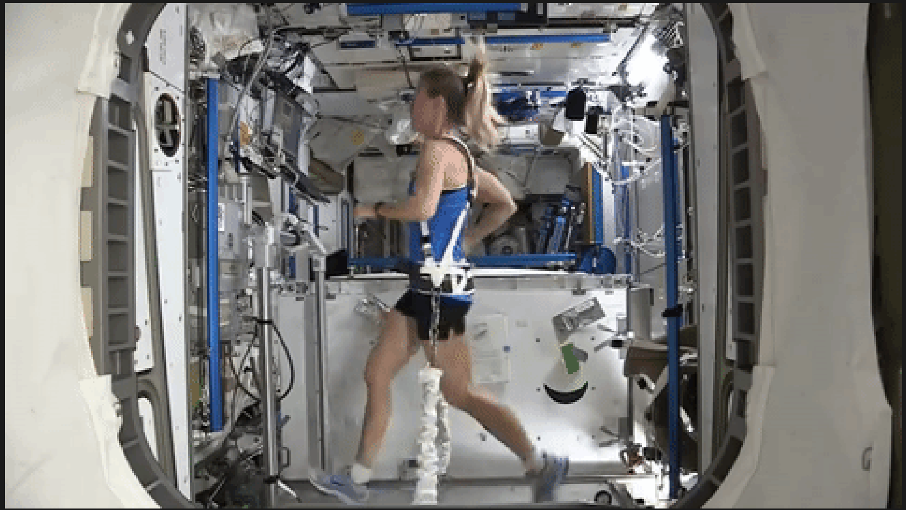
- Software automatically compensates for camera pitch angle



Intelligent Personal Coach - Results

Intelligent Personal Coach - Future work

- Create visualizations for other exercises
- Utilize more robust UI framework for visualization
- Gamification
- Hardware optimization



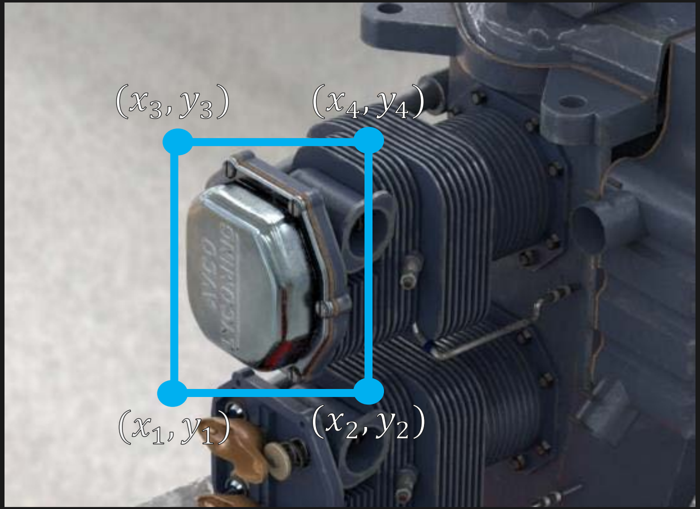
Safety Analysis with Deep Learning

DML Safety Analysis - Motivation

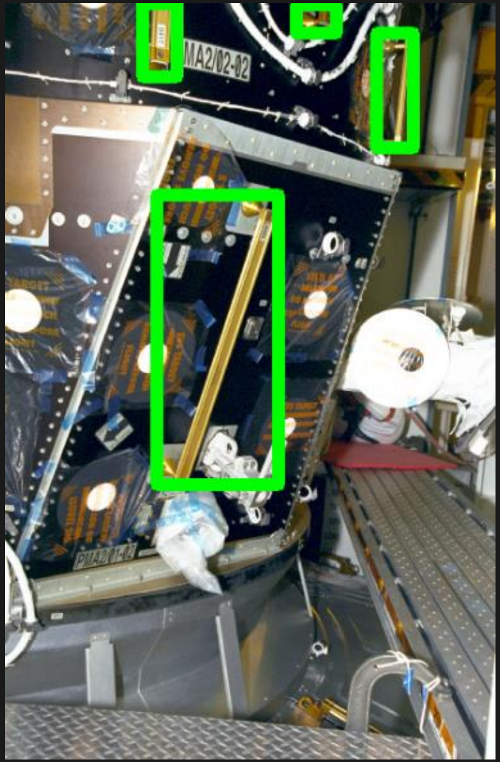
- Micrometeoroid and Orbital Debris (MMOD) damage identification on the ISS:
  - MMODs on handrails can tear spacesuit gloves.
  - Tedious process: human in front of the screen zooming and inspecting images manually for hours
  - Introduces opportunity for human error due to fatigue, attention span, etc.
- Develop a system to identify handrails in images
  - Goal: assist imagery team by automating this manual and laborious process.

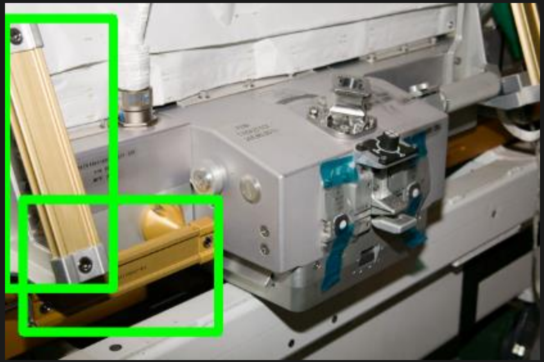
Faster R-CNN - 2D Case

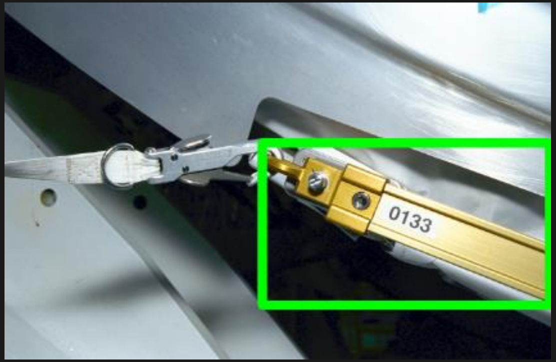
- Trains to predict bounding boxes around image objects
  - 2 steps: localization, and classification
- Training data: both images and bounding box coordinates
- Network will predict 4 points in pixel space and a class for each object it finds

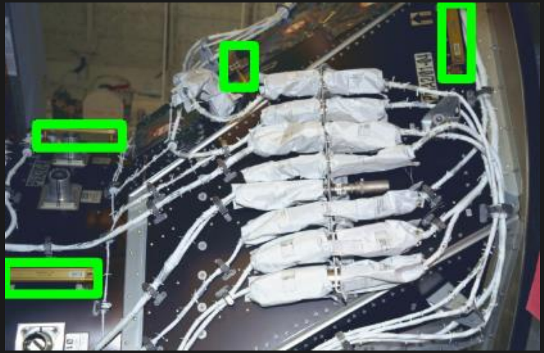


DML Safety Analysis – Training Pictures

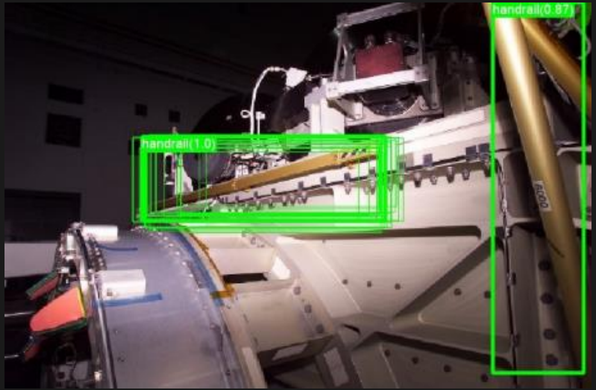


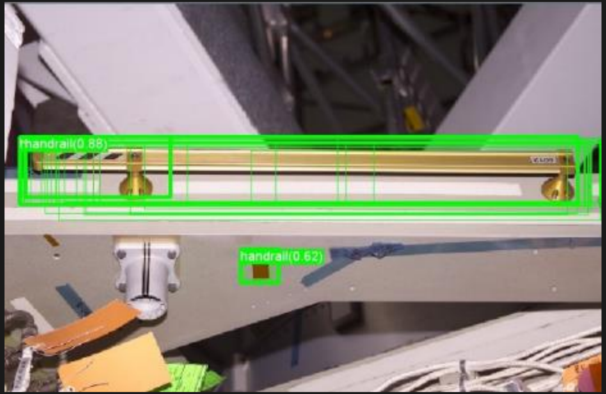


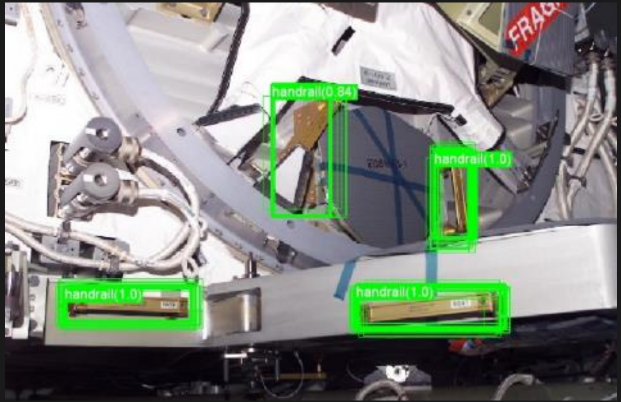


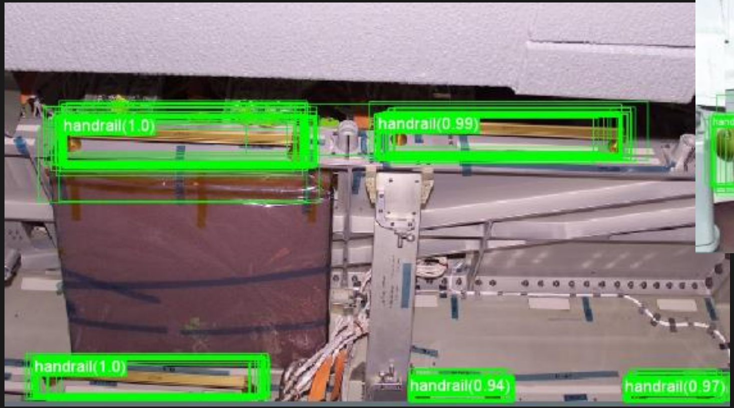


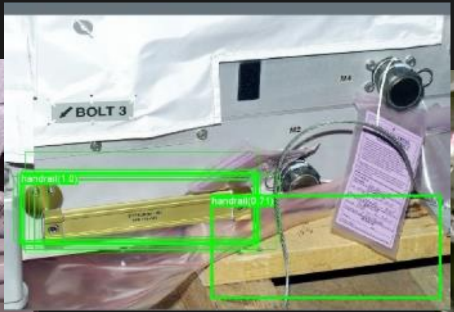
DML Safety Analysis – False Positives

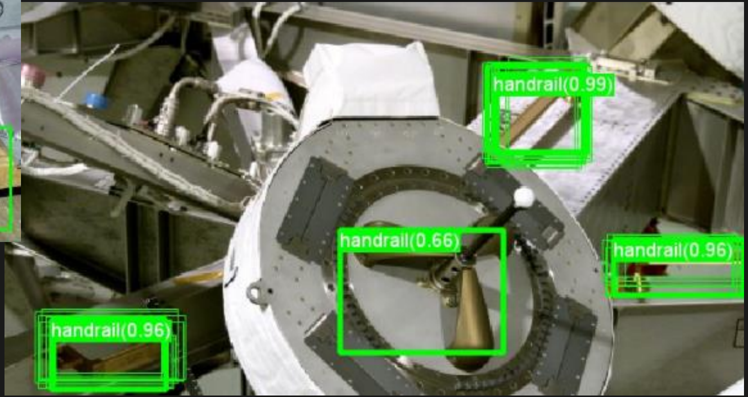




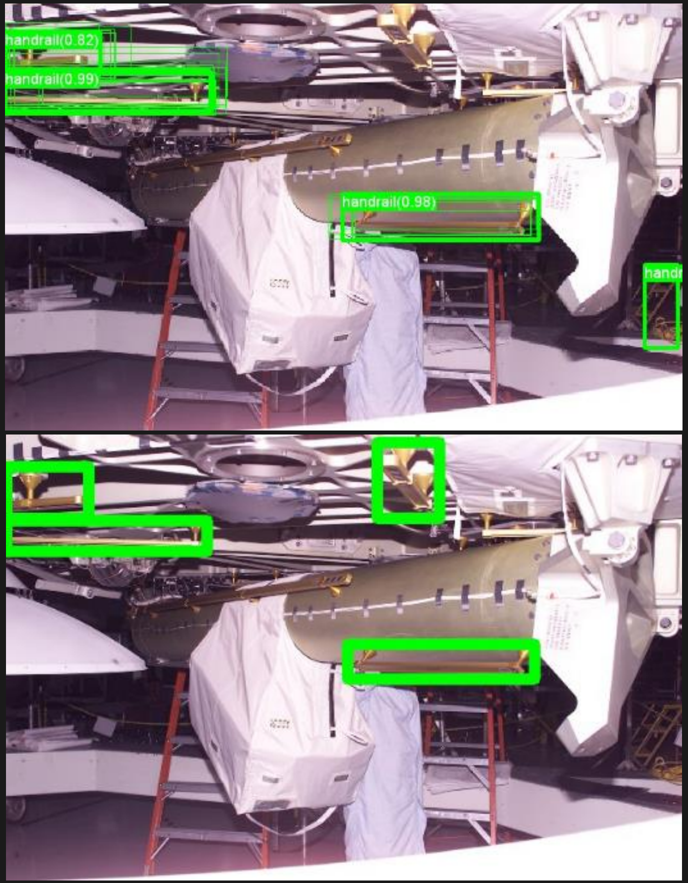


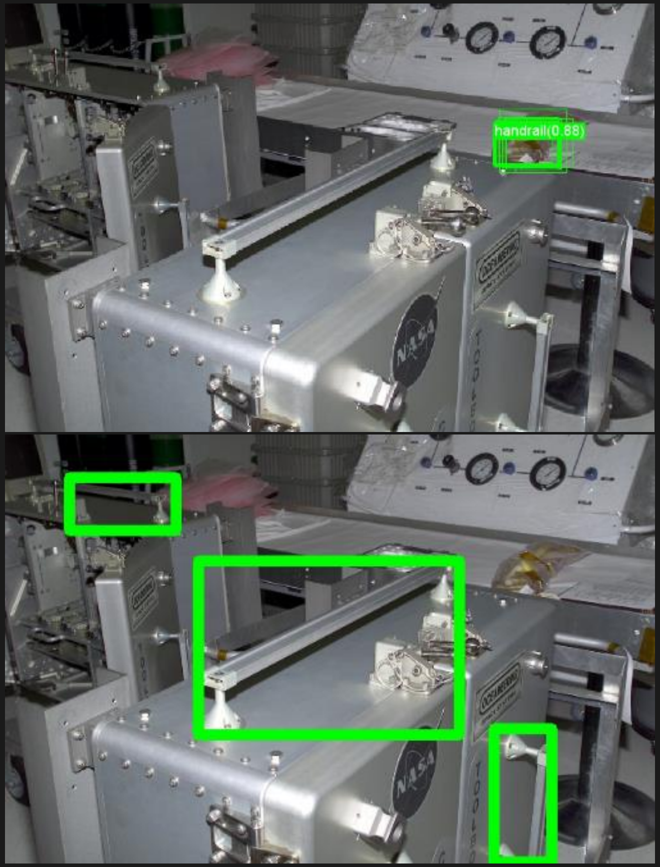






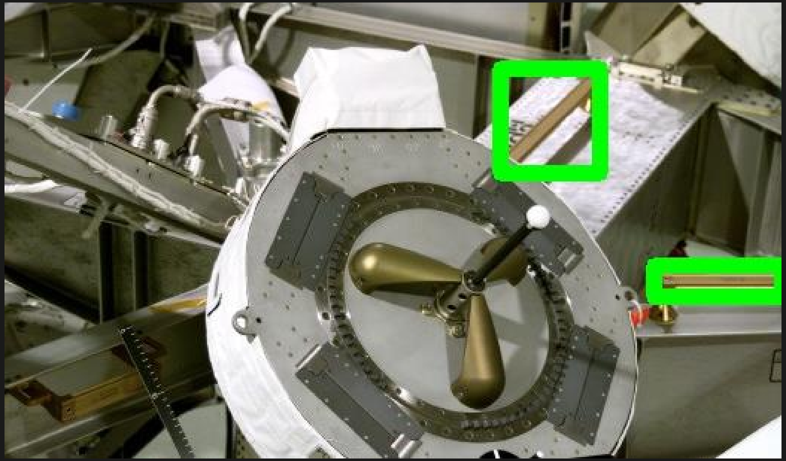
DML Safety Analysis – False Negatives





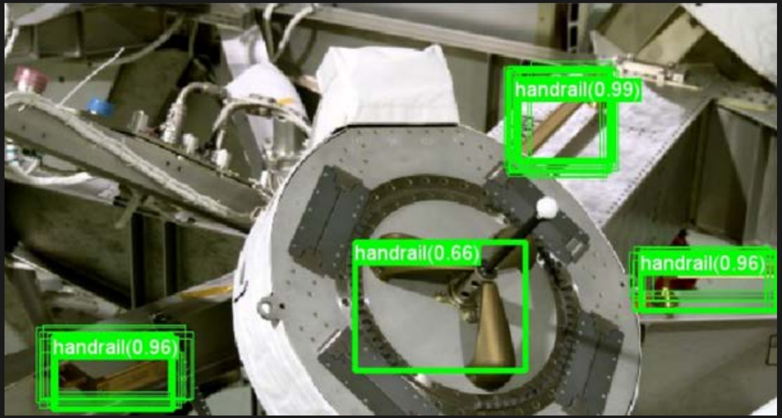
Better and Worse than Human

Human labeled image

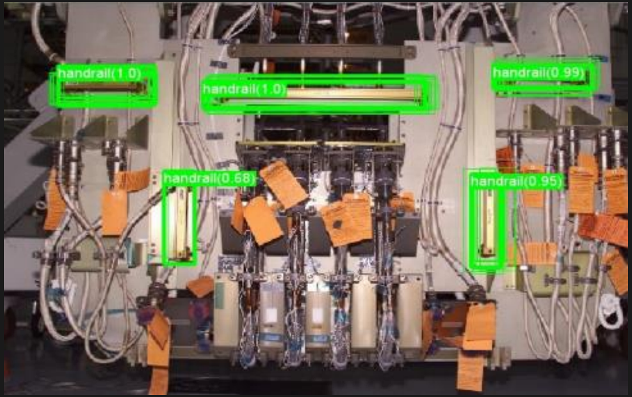


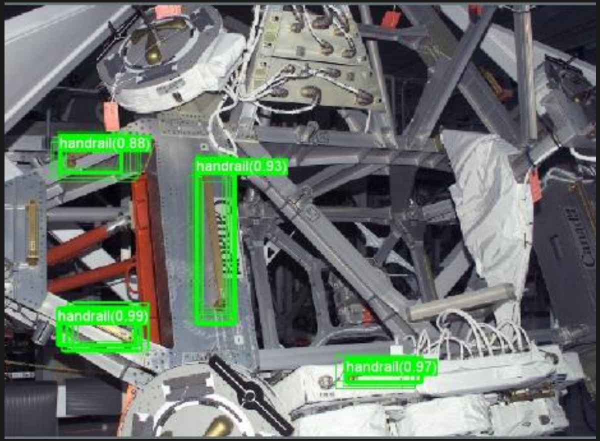
DML resulting labeled image:

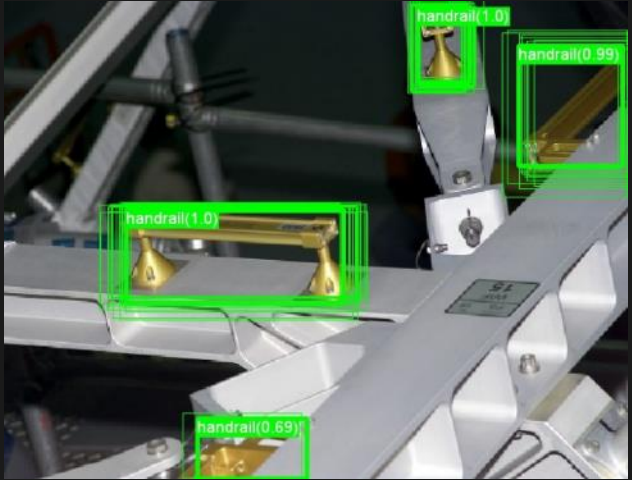
- Found a missing handrail from human labeled image
- Found a False Positive

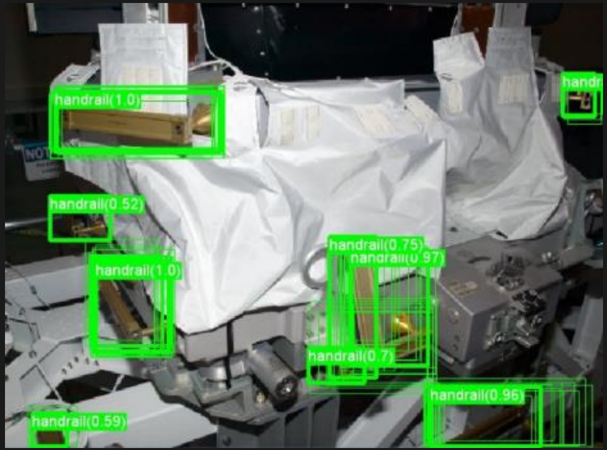


Success in Complex Images

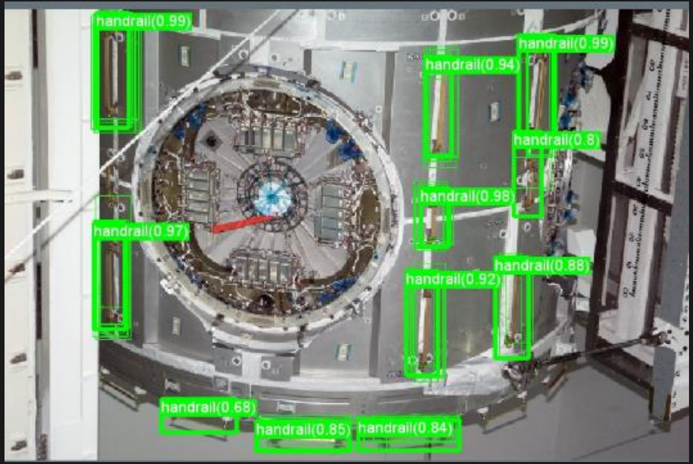


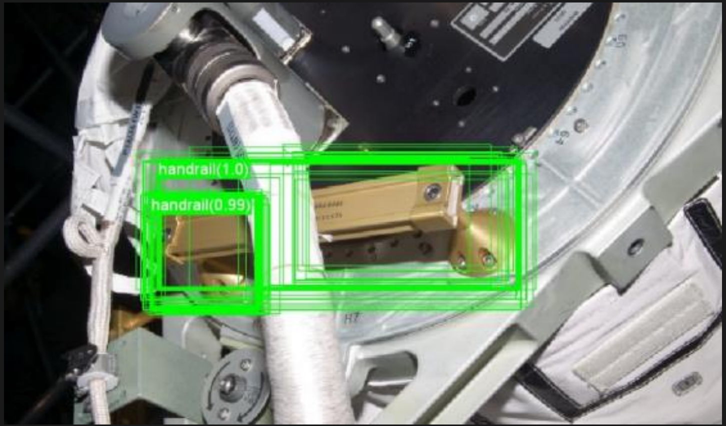


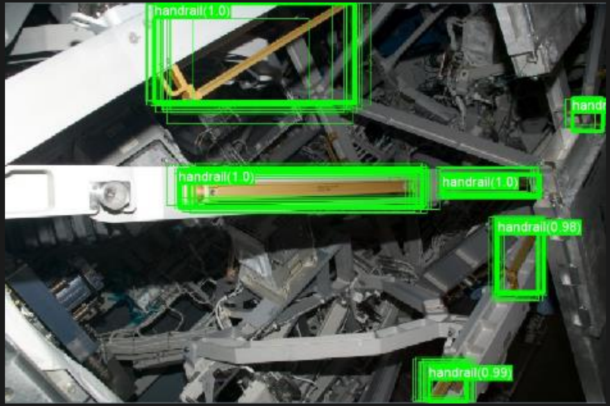


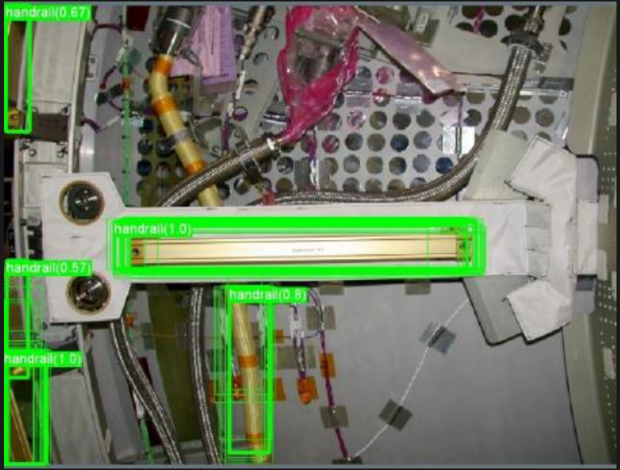


Success in Complex Images









DML Safety Analysis - Issues

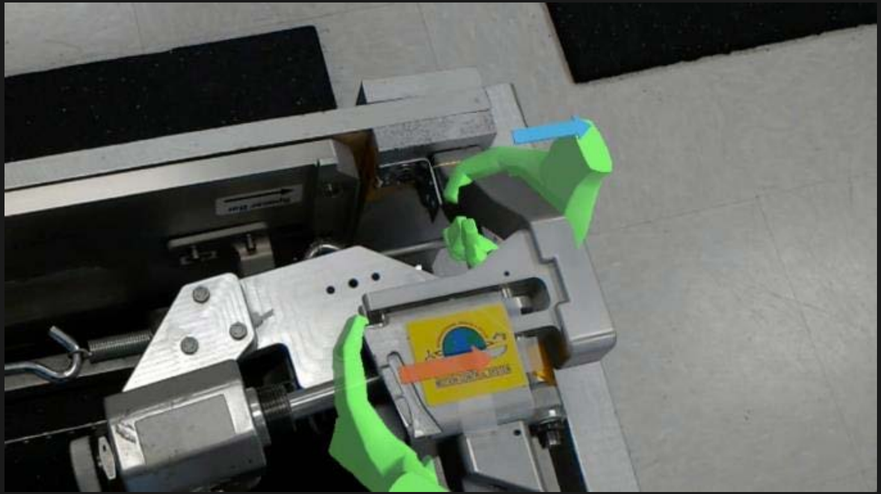
- Number of images for training and validation too low
- Existing labels of images not accurate
- Labeling images is manual intensive

6DOF Object Pose Estimation with Virtual Training Dataset

Object Pose Estimation - Background

- Motivation - AR Object Alignment
  - AR procedure assistant needs to align a 3D model with reality
  - Traditionally done with sticker anchors

- Main Problems
  - If the object moves relative to the sticker, the AR model cannot compensate.
  - Can only register by looking directly at the sticker
  - Inertial drift



Object Pose Estimation - Background

- Goal - eliminate the need for stickers using deep-learning based pose estimation.
    - “Continuous” object registration to eliminate drift
    - Support for partial occlusion, many more viewpoints
    - Support for different geometries

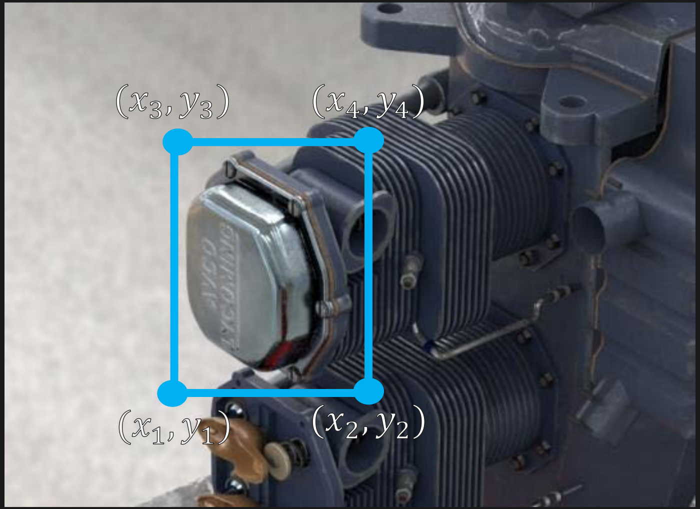
- Requirements
    - NN that takes RGB image as input, produces 6-DOF pose estimation as output
    - **Large, diverse training dataset of 6-DOF labeled images.**

Object Pose Estimation - Approach

- **Acquiring and manually labeling a large enough dataset to be useful is a near-impossible task**
- Instead, use photorealistic renderings from CAD model as training data
  - Generate thousands of pre-labeled images quickly
  - Easily vary parameters (background, clutter, distortion, lighting, etc)
- Two problem phases:
  - Test viability of synthetic training data on simplified 2D case
  - Advance to full 6-DOF predictions
- Can we generate synthetic images with high enough quality to train on for a real test set?

Object Pose Estimation - 2D Case

- Detecting piston covers on Ellington Field airplane engines:
  - 3k-5k synthetic images rendered for each iteration
  - Bounding boxes automatically generated in Maya
  - ~50 photographs and several videos for testing
  - Trained TensorFlow implementation of **Faster R-CNN**
  - Relatively rapid series of small scale tests

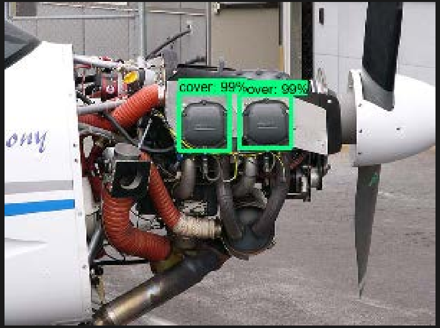


Object Pose Estimation - 2D Case

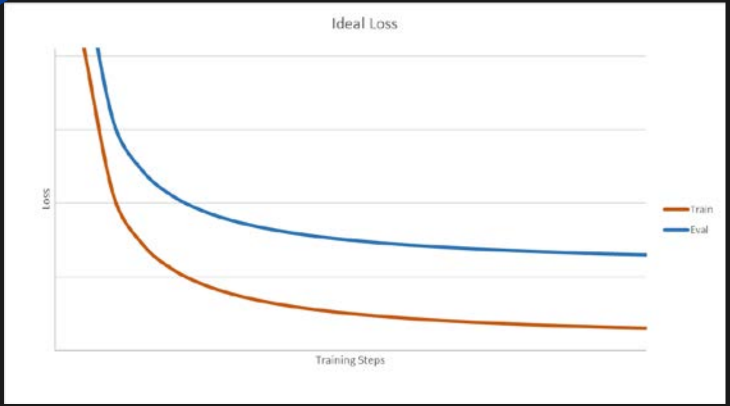
- Identified sensitivities of network through many training iterations
- Gradually added complexity to dataset (clutter, colors, reflectiveness, lighting, etc..)
- Gradually tuned hyperparameters of network (batch size, learning rate, data preprocessing etc..)
- Eventually reached >90% accuracy

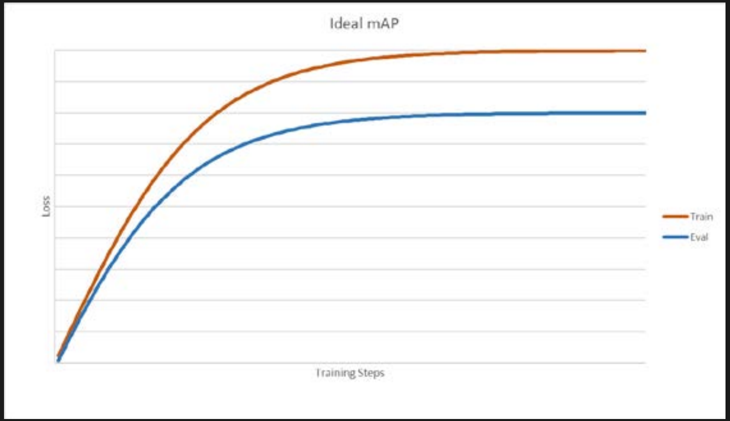
Object Pose Estimation - 2D case

© Red Bull Air Race / https://www.youtube.com/watch?v=OtdiLc_wajg



Implications – Ideal Case

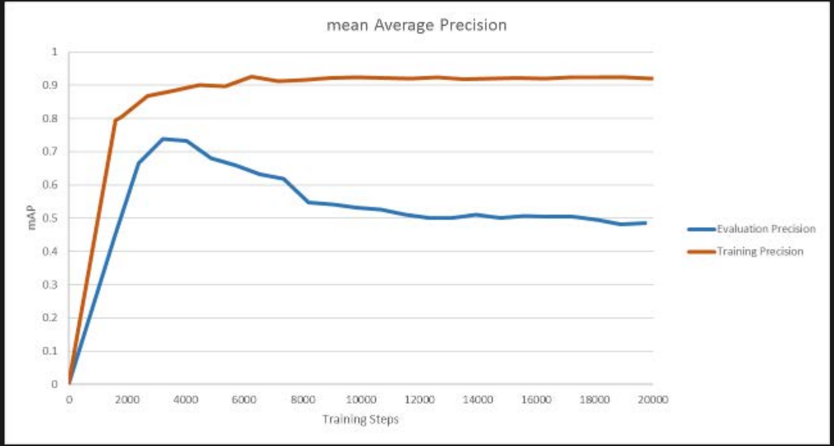




- **Key observations:**
- Steep at first, then follows a shallow gradient
- Both sets have the same general shape, even though the eval set has worse performance

- **Our configuration:**
- 1000 synthetic training images
- ~30 manually labeled test photographs

Our Results



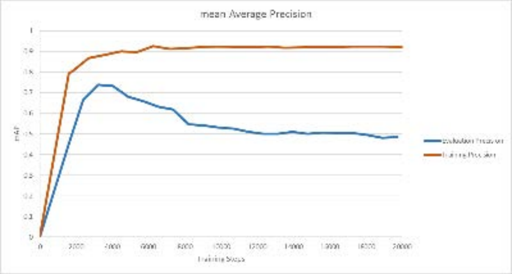
Key observations:

- Evaluation set reaches optimal values at ~2k steps
- Training set performance converges quickly, while evaluation set diverges

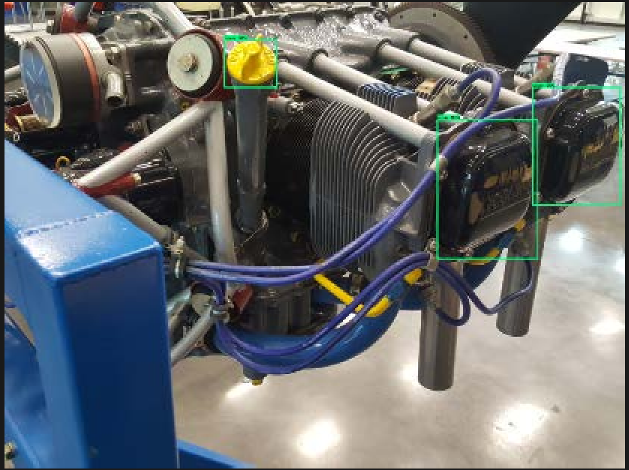
What does this mean?

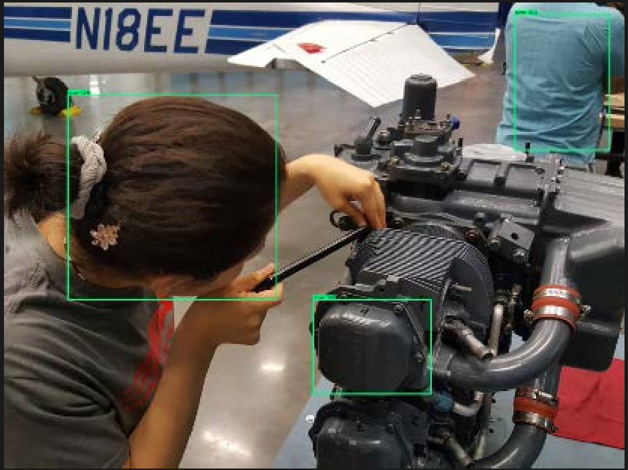
- Model is quickly *overfitting* to training set

Our Results



Step 20000



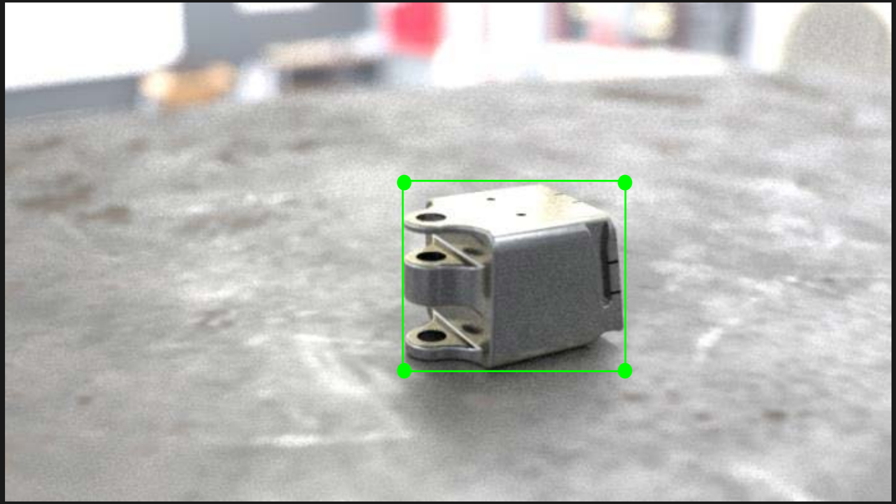


Object Pose Estimation - 3D case

- Larger training set
  - ~10k images or more
- More diverse training set
  - Lighting
  - Materials
  - Background
  - Occlusion
  - Color
  - Focus
  - **More negative examples**
- Other considerations
  - More photos for evaluation
  - Trying different algorithms and cross training
  - Explore data augmentation with random filters, distortion and scaling

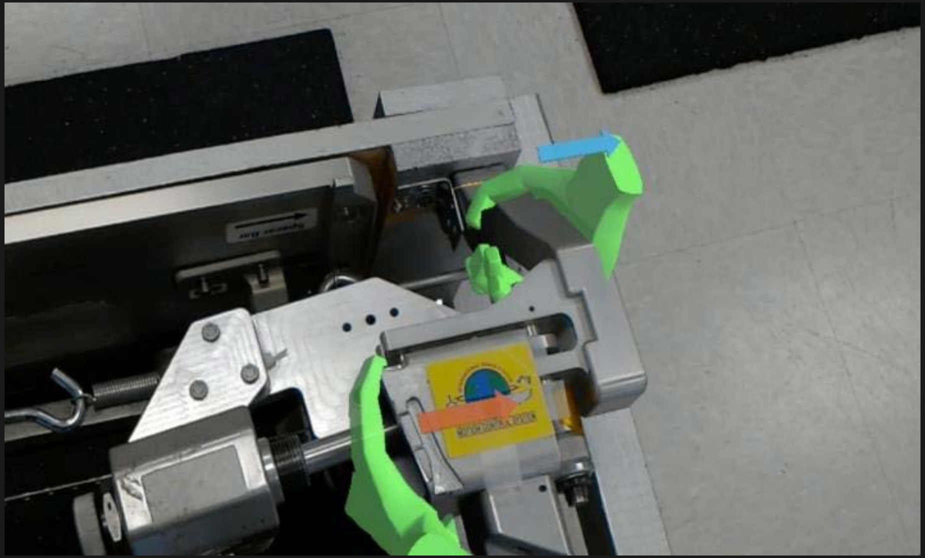
Object Pose Estimation - 3D case

- 6-DOF pose estimation is a much harder problem than 2D bounding boxes
- Singleshotpose - new algorithm developed by Microsoft
  - Based on the same principle as 2D bounding box estimation
  - Instead, train to predict 2D *projection* of 3D bounding box
- Divides the image into a series of boxes, with identified centroids in a box, along with point offsets relative to the centroid (YOLOv2).



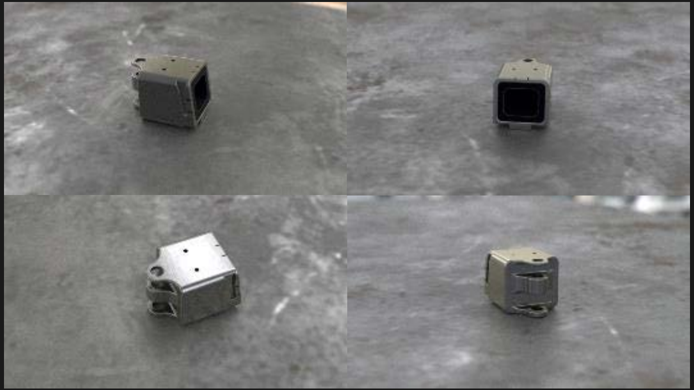
Object Pose Estimation - PnP

- Use PnP algorithm to recreate 3D coordinates with respect to camera
- 3D points can be used in AR procedure assistants

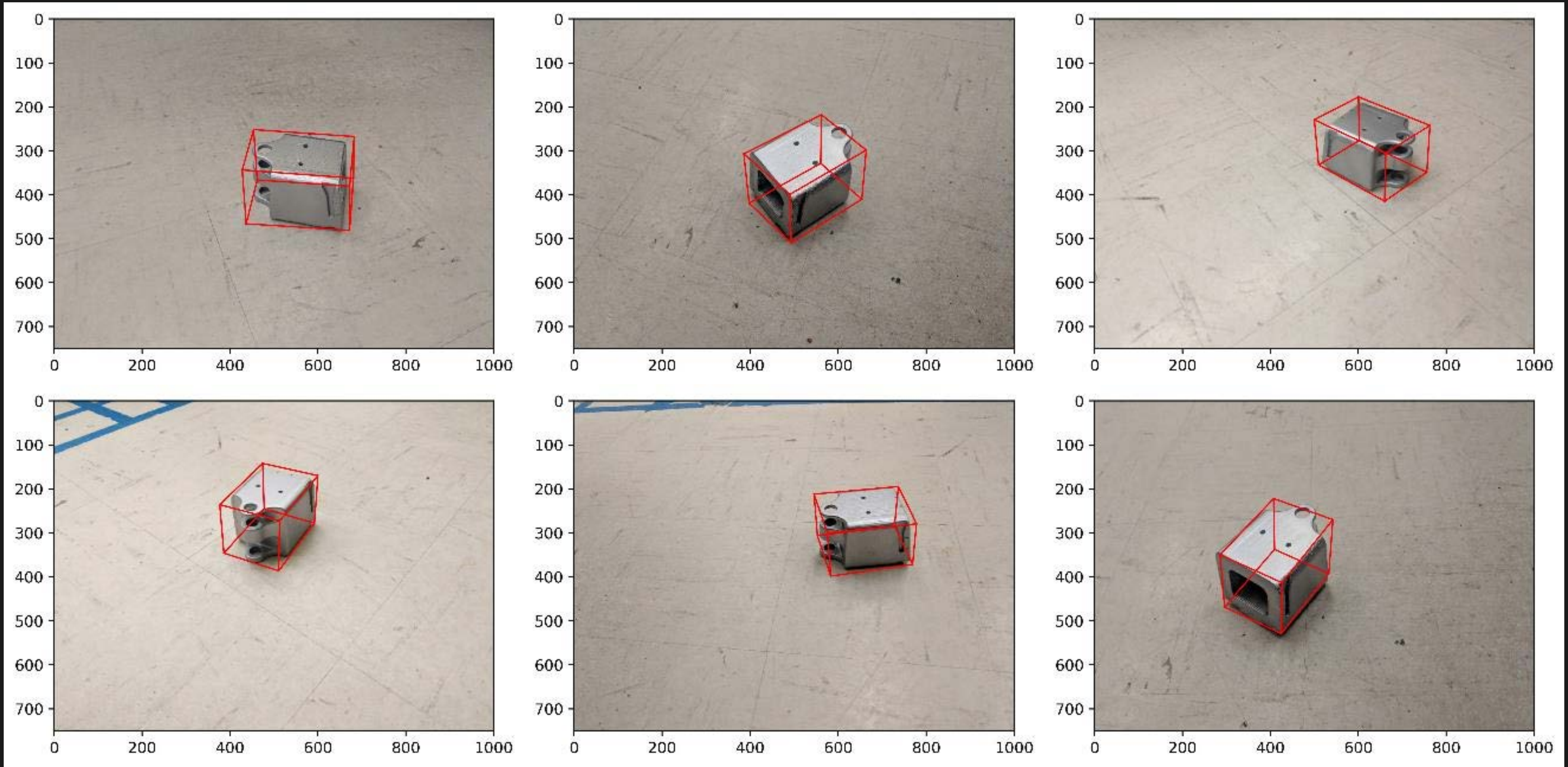


Object Pose Estimation - 3D case

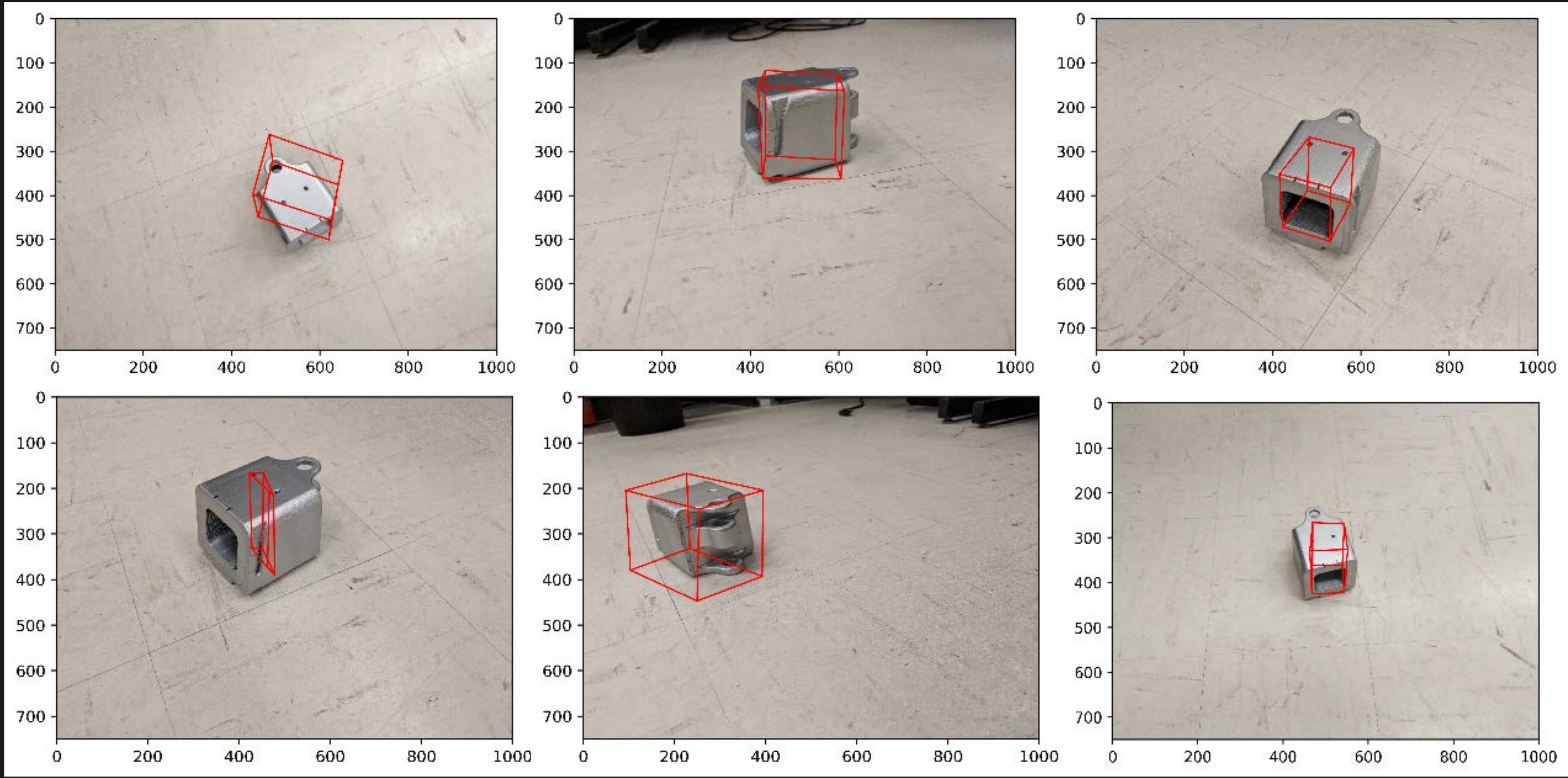
- Snubber-cup pose estimation:
  - 3k synthetic images rendered
  - 3D bounding boxes and 2D projections automatically generated with Maya
  - 3D printed model for testing
  - Trained slightly modified reference implementation of singleshotpose



Object Pose Estimation – Best Results



Object Pose Estimation – Worst Results…



Object Pose Estimation - Future Work

- Provide clearly defined metric for accuracy
  - Planned out workflow for using stickers as ground-truth for future iterations
- Gradually add complexity to system
  - Realistic background, lighting, partial occlusion, camera noise, etc.
- Create pipeline for future use cases
  - CAD model as input - trained network as output

Bigger Picture - DML at NASA

- Biometric Monitoring
  - Pupillometry, EEG, fNIR
  - Predict physical and mental health problems
- Vehicle Systems Management
  - Augmented Reality Procedure Assistants
  - Autonomous Surface Vehicles and Navigation
  - Mission Control AI

Acknowledgements

- **Lawrence Hessburg** – Graduate Student
- **Lui Wang** – Augmented Reality Expert
- **Frank Delgado** – Advanced Operations Concepts Lab Manager

Questions