Coursework 2 – COMP0248

Dataset

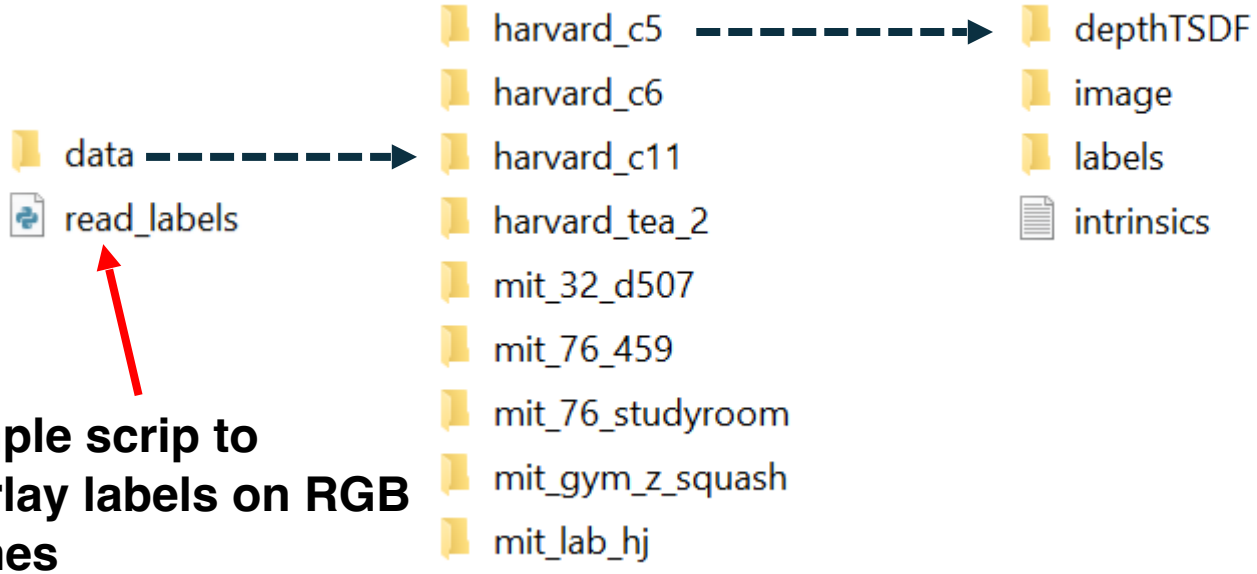
- Selected data from Sun 3D

Depth

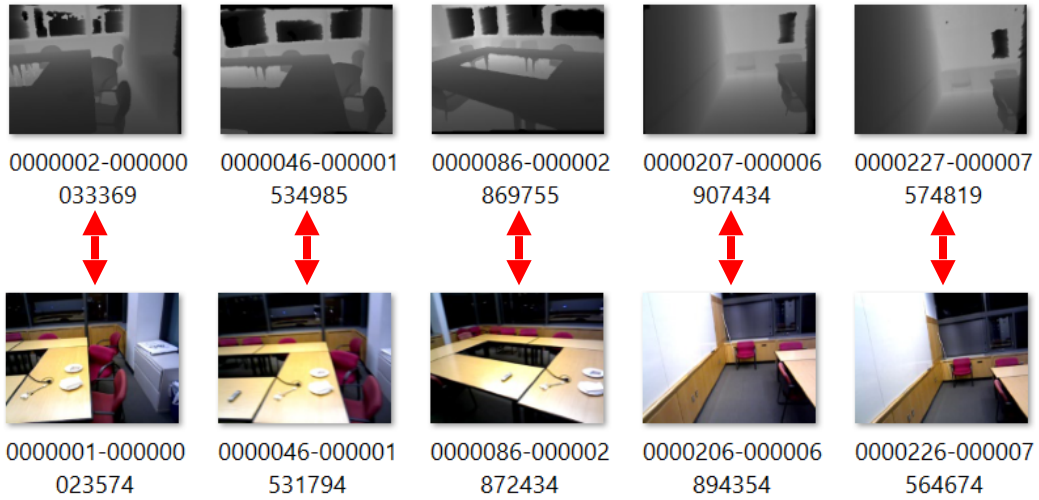
RGB

Table Polygon Annotations

Dataset



Sample scrip to overlay labels on RGB frames



File name: `frame-timestamp`

Depth and RGB frames are paired by closest timestamp, not frame number (files are already in the correct order).

Tasks: Binary classification (Is there a Table in the image?)

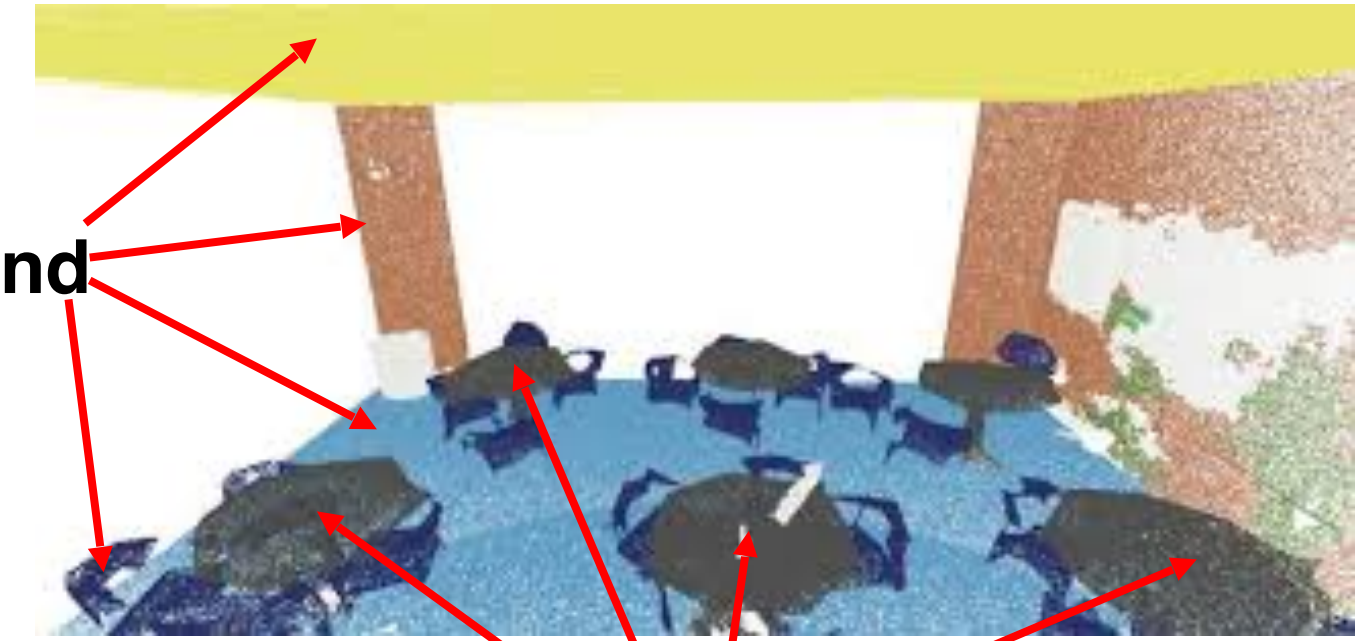
Table

No Table

- **The following labels from Sun3D were grouped under a single "table" class: "`table top`", "`dining table`", "`desk`", "`coffee table`"**
- **Does not include: cabinets, kitchen counter**

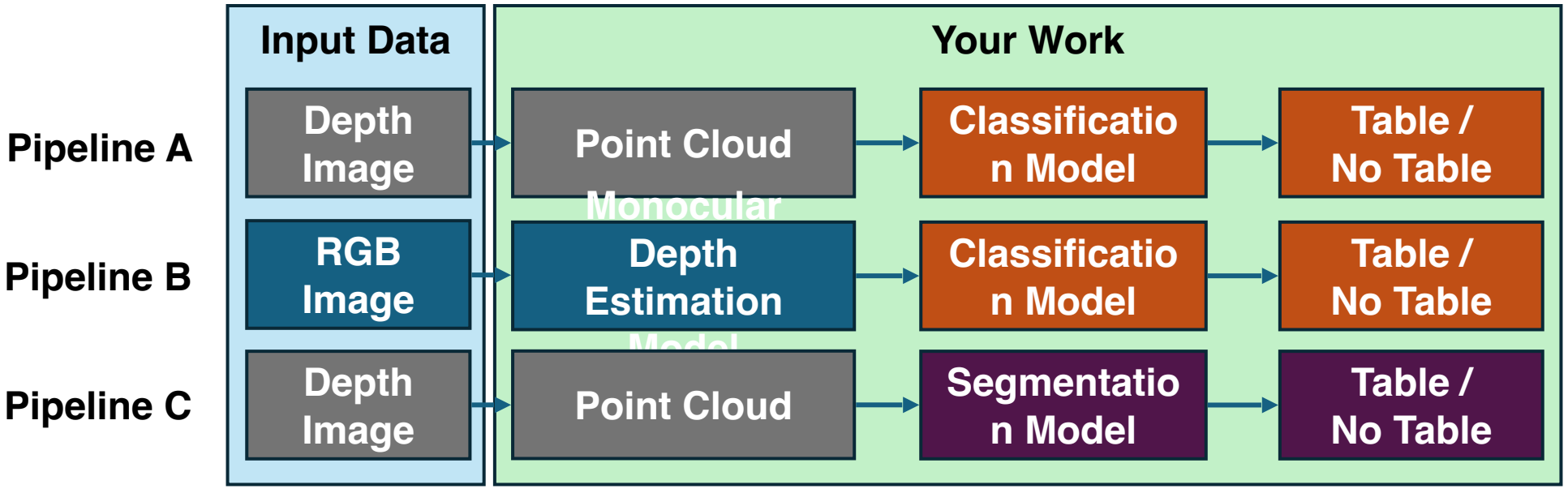
Tasks: Binary Point Cloud Segmentation

- Classify each point as Table or Background



Implement 3 different pipelines

- **Pipeline A:** Convert a depth input to a point cloud, then a classifier takes a point cloud as input
- **Pipeline B:** First a model estimates depth from a 2D RGB image, then a classifier takes the estimated depth map as input
- **Pipeline C:** Convert a depth input to point cloud, then a segmentation model takes a point cloud as input



Model Implementation and Training

- **Overall:**
  - **You can re-use/import existing models, without implementing their architecture from scratch.**
  - **You need to explicitly state and cite the works you are using in your report.**
  - **For full marks, use models different from tutorial examples (vanilla PointNet, MonoDepth2 pre-trained on KITTI, etc).**
- **Monocular Depth:**
  - **You can use off-the-shelf pre-trained models if they work reasonably well on the test data (no penalty on marks).**
  - **You will need to explicitly state and cite the works you are using in your report.**
- **Classification/Segmentation:**
  - **You can use pre-trained models as initialisation (cite on report), but you need to train them on the CW2 dataset**

Coursework2 Dataset - Data Split

- **Training Data:**
  - **MIT sequences**
  - `mit_32_d507, mit_76_459, mit_76_studyroom, mit_gym_z_squash, mit_lab_hj`
  - **Total 290 RGBD frames**

- **Test Data 1:**
  - **Harvard sequences**
  - `harvard_c5, harvard_c6, harvard_c11, harvard_tea_2`
  - **Total 98 RGBD frames**

- **Test Data 2:**
  - **UCL sequence: capture your test data with an Intel RealSense Camera**
  - **Select a maximum of 50 RGBD frames from any video captures**
  - **Quantitative classification/depth evaluation**
  - **Qualitative segmentation evaluation (no need to annotate point clouds)**

Other dataset details

- **Sequences with no labels / no tables (negative samples):**
  - `mit_gym_z_squash`
  - `harvard_tea_2`
- **Detected missing table labels:**
  - `76-1studyroom2 - 0002111-000070763319.jpg`
  - `mit_32_d507 - 0004646-000155745519.jpg`
  - `harvard_c11 - 0000006-000000187873.jpg`
  - `mit_lab_hj - 0001106-000044777376.jpg`
  - `mit_lab_hj - 0001326-000053659116.jpg`
- **Depth maps have been pre-processed by DepthTSDF for improved quality, except for `harvard_tea_2`, which contains raw depth maps from camera. Raw data acquired with RealSense will be more similar to `harvard_tea_2` data than the other sequences**

Report Sections [maximum 6 pages + references]

- **Introduction / Problem Statement [~0.5 pages]**
  - **Precisely define the problem you are solving, detailing any assumptions you are making about the data when designing your models**
  - **Cite the works you use as inspiration for your models, explaining why you chose them (pre-trained weights, architecture, training methodology, etc)**
- **Data processing [~1 page]**
  - **Describe how you processed data to generate groundtruth classification/ segmentation labels from polygon labels**
  - **Describe any pre-processing steps**
  - **Describe how you convert between different data types (e. g. from depth to point clouds, etc)**
- **Methods [~2 pages]**
  - **Pipeline A**
  - **Pipeline B**
  - **Pipeline C**

Report Sections [maximum 6 pages + references]

- **Results [~1.5 pages]**
  - **Classification performance for Pipeline A, Pipeline B**
  - **Depth Estimation performance for Pipeline B**
  - **Segmentation performance for Pipeline C**
  - **Describe and justify the metrics you use and any other evaluation related processing steps**
- **Discussion [~0.75 pages]**
  - **Compare performance of Pipelines A/B, describe and justify relative strengths/weaknesses based on obtained results**
  - **Identify Pipeline C strengths and weaknesses based on obtained results**
- **Conclusions [~0.25 pages]**
  - **What are the limitations of the explored approaches?**
  - **How could the methods be improved?**
- **References [use as much space as you need]**

Marking [total 100]

- [16]: Good understanding of the problem
  - Report: “Introduction”, “Discussion”, “Conclusions”
- [16]: Data processing
  - Code
  - Report: “Data Processing”
- [36]: Model Implementation
  - Code
  - Report: “Methods”
  - [12] for each pipeline
  - Up to [8] for each pipeline if you use exact same model as tutorials (vanilla PointNet, MonoDepth2 pre-trained on KITTI, etc)
- [20]: Evaluation on CW2 dataset
  - Code
  - Report: “Results”
  - Maximum [10] if only 1 pipeline is evaluated
  - Maximum [15] if only 2 pipelines are evaluated
- [12]: Capture and evaluation on RealSense data
  - Code
  - Report: “Results”
- Except for references, anything beyond the page limit (6 pages) will be ignored for marking

Marking – General guidelines

- **Justify your implementation choices (what favours them compared to alternatives?)**
- **There is no strict minimum model performance you need to achieve. For results, focus on rigorous, fair, and well justified evaluation. If your model fails sometimes, explain what do you think the reasons are in the discussion section.**
- **Use figures/diagrams if it makes explanations clearer**
- **Clearly label plots and tables.**
- **Clearly define symbols/variables if you write equations**
- **For discussion/conclusions, avoid generic statements. Write based on your own obtained results.**

Marking – General guidelines

- **Equal marking for all group members, unless exceptional circumstances (e. g. a group member did not work and did not contribute to the coursework submission).**
- **If you have any issues during coursework period, communicate with staff (Sophia, Francisco) as soon as possible.**

Submitting your report

- **Use the latex template IEEE Conference template**
- **https://www.overleaf.com/latex/templates/ieee-conference-template/grfzhhncsfqn**

Submitting your code

```
Code/
├── src/
│   ├── pipelineA/    # organise how you want, but
│   ├── pipelineB/    # keep a modular structure and
│   └── pipelineC/    # clearly explain in README.md
├── data/
│   ├── CW2-Dataset   # delete for submission
│   └── RealSense     # include in submission
├── results/          # predictions, logs, plots
├── weights/          # delete in submission
├── requirements.txt
└── README.md
coursework2_groupXX.pdf
```