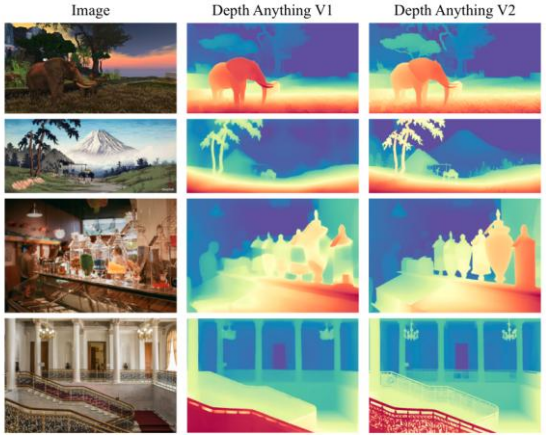
TECHNISCHE
UNIVERSITÄT
DARMSTADT

Potential Research Topics

Prior-Guided Dense Depth Estimation for RGB Images

TECHNISCHE UNIVERSITÄT DARMSTADT

- In robotics it is important to have 3D understanding of the world, not just on the image plane
- to transform 2D detections to 3D detections, we often rely on RGB → LiDAR mapping, which can be sparse
- It would be best if we have a direct mapping pixel → 3D coordinate (i.e. a depth map for the image)
- There exist many networks that yield this representation (i.e. depth Anything), however they are not metric (i.e. only relative values)
- Recently there is emergent work in prior guided metric depth estimation that allows for metric depth estimation



Prior-Guided Dense Depth Estimation for RGB Images

- State-of-the-Art:
  - [1] Lin, H. et. al. „*Prompting Depth Anything for 4K Resolution Accurate Metric Depth Estimation*“. arXiv December 18, 2024. https://doi.org/10.48550/arXiv.2412.14015.
  - [2] Wang, Z. et. al. *„Depth Anything with Any Prior“.* arXiv May 15, 2025 https://arxiv.org/abs/2505.10565
- Research Gaps:
  - the effect of different sampling strategies (different LiDARs = different patterns)
  - real-time capability
  - uncertainty quantification

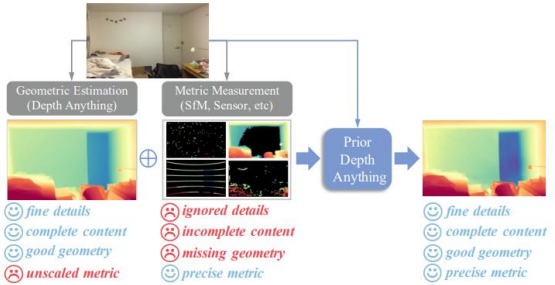


Figure 1. **Core Motivation**. We progressively integrate complementary information from metric measurements (accurate metrics) and relative predictions (completeness and fine details) to produce dense and fine-grained metric depth maps.

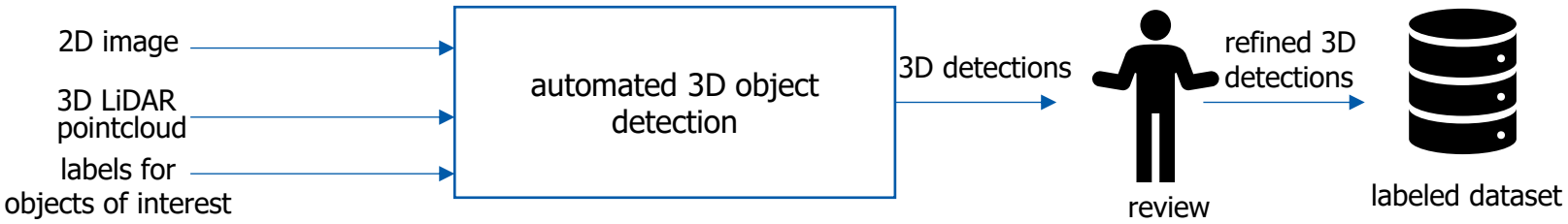
Prior-Guided Dense Depth Estimation for RGB Images

TECHNISCHE UNIVERSITÄT DARMSTADT

- Potential work packages:
  - Initial testing of existing work
  - Capturing of indoor/ outdoor dataset with different LiDARs and RGB-D cameras
  - Investigate effect of different sampling patterns
  - Optional, but nice contributions:
    - Adaptation/ distillation of network for realtime use
    - uncertainty quantification

Open-World 3D Object Detection Pipeline

- To train 3D LiDAR-based object detectors, we need labelled data of the objects of interest
- Manual annotation is time-consuming
- Having a pipeline for automated 3D object detections that just needs review (i.e. in CVAT) would greatly reduce the effort needed
- This needs effective algorithms for 3D cluster proposal generation



Open-World 3D Object Detection Pipeline

- State-of-the-Art often uses Early-Fusion Approaches

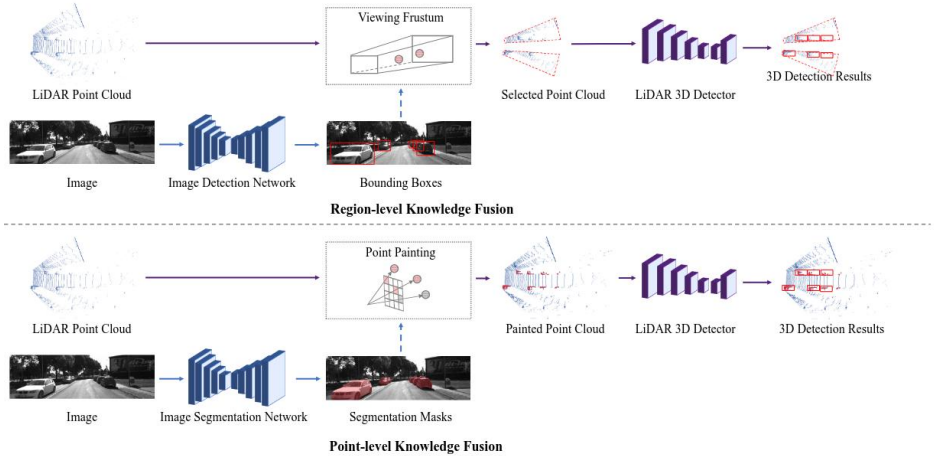


Fig. 17: An illustration of early-fusion based 3D object detection methods.

Mao et. al. "3D Object Detection for Autonomous Driving: A Comprehensive Survey." arXiv, April 4, 2023

Open-World 3D Object Detection Pipeline

- example Pipeline from State-of-the-Art:
  - Xia, Z. et. Al. *„OpenAD: Open-World Autonomous Driving Benchmark for 3D Object Detection”*, arxiv, November 26, 2024

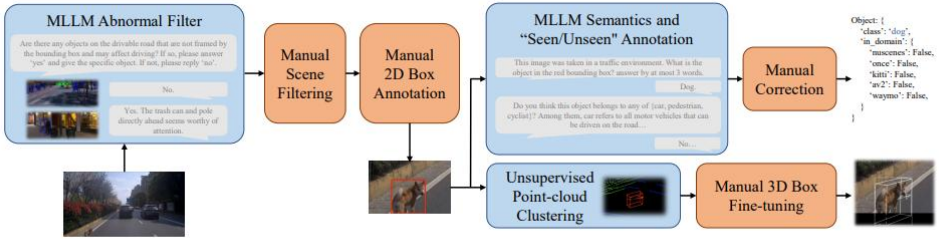
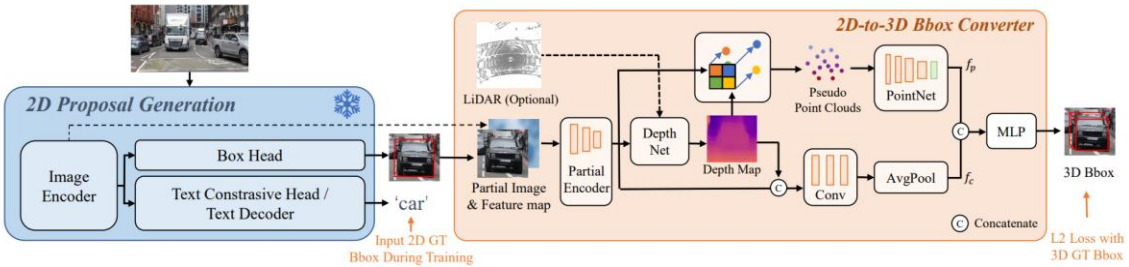


Figure 3: **Annotation pipeline**. OpenAD is built upon a corner case discovery and annotation pipeline that integrates with a multimodal large language model (MLLM).



Open-World 3D Object Detection Pipeline

TECHNISCHE UNIVERSITÄT DARMSTADT

- Potential Work Packages:
  - Literature Research to existing LiDAR clustering techniques
  - Selection of promising approaches
  - Capturing of dataset
  - Investigation of effectiveness of different clustering techniques
  - Integration into CVAT labeling process