Smart City Road Maintenance: A LiDAR and AI-Driven Approach for Detecting and Mapping Road Defects

Authors: Giovanni Nardini*, Roberto Nucera*, Alessandro Ulleri*, Stefano Cordiner†, Eugenio Martinelli†, Arianna Mencattini†, Iulian Gabriel Coltea*

*Key to Business, Italy, †University of Rome Tor Vergata, Italy

Presenter: Giovanni Nardini, R&D AI Tech Lead, Key2, info@key2.it

Valencia, Spain April 06-10, 2025

Giovanni Nardini

Giovanni Nardini holds a Master's degree in Mechanical Engineering from Sapienza University of Rome, and specialized in AI and Computer Vision. His current work involves combining insights from his engineering background with Artificial Intelligence and Computer Vision, focusing on developing solutions for real-world challenges. He now leads AI development at Key2, focusing on Computer Vision and Machine Learning applications.

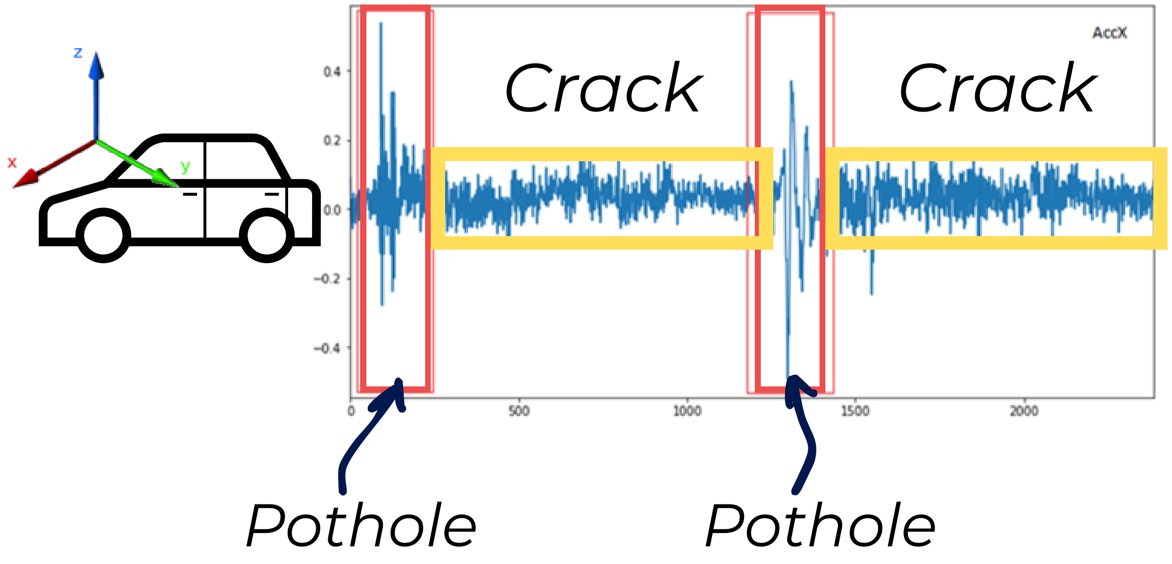
The challenge of maintaining roads

Traditional Road Maintenance

- Lot of **people** involved
- **Unefficient** scheduling of reparations
- **Unsafe**: no rapid response to road issues
- **Sparse** and incomplete data
- Very **Expansive**
- Suffers form subjectivity & **inconsistency**
- **Manual** inspection and **slow** processes

Alternative technological approaches ..

In-Vehicle Inertial Data Analysis:

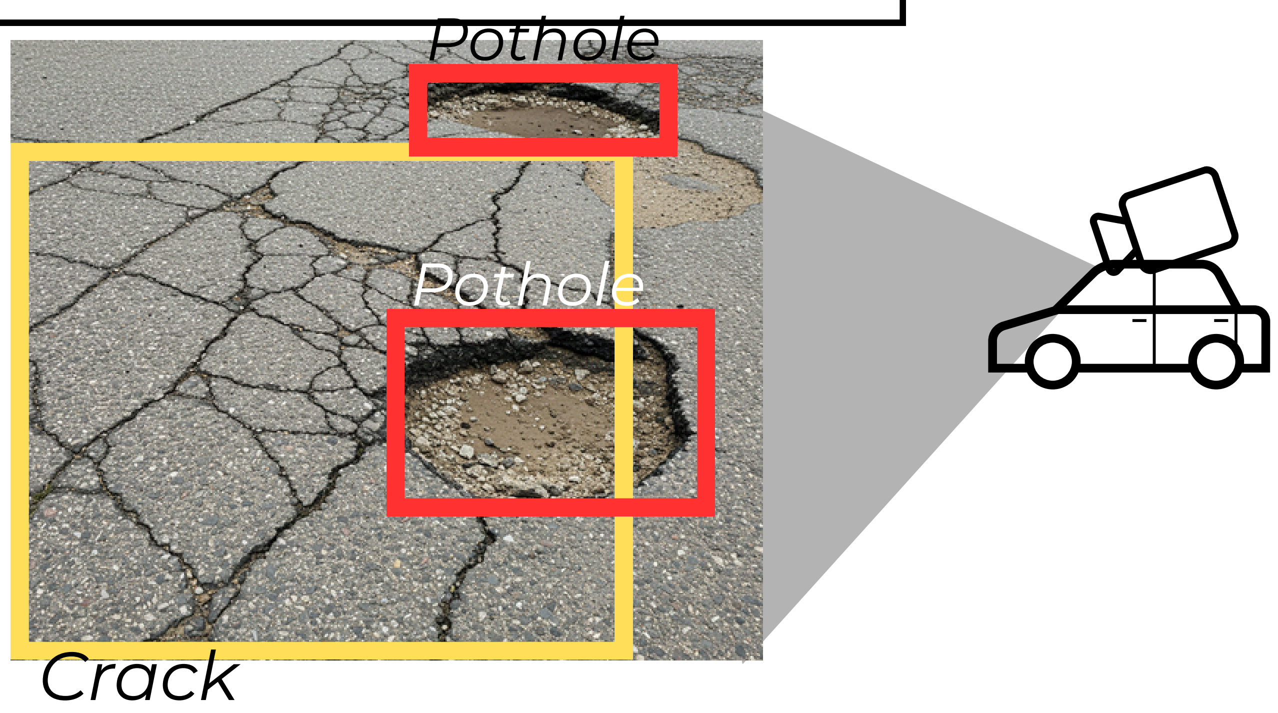


Can detect impacts and noise

BUT:

- limited detail on geometry
- depends on defect size
- depends on vehicle path
- Noise interpretation

Computer Vision & Deep Learning on camera:



Can see defects in images

BUT:

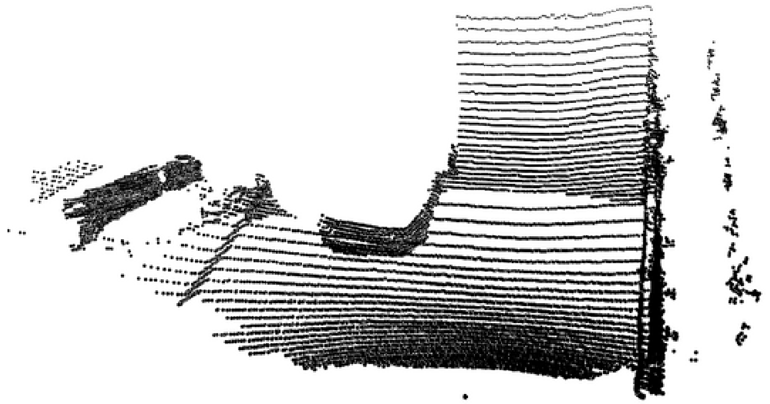
- location accuracy is poor
- suffers from lightning conditions
- detections are in 2D, very poor size estimation

Our Approach

USB CAMERA

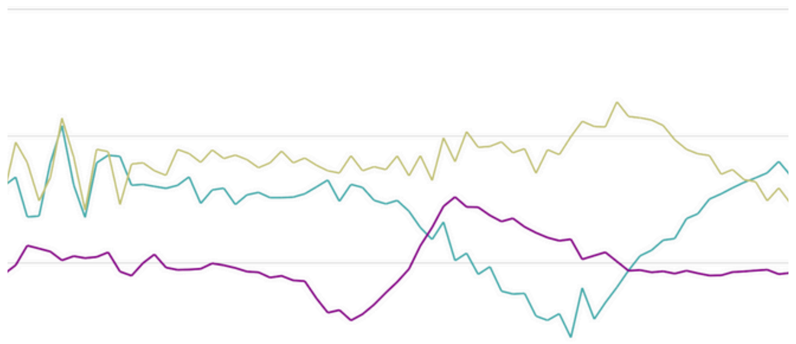
Vision Data: RGB Image

128 Channels LiDAR



3D Point Cloud Data

IMU/GNSS



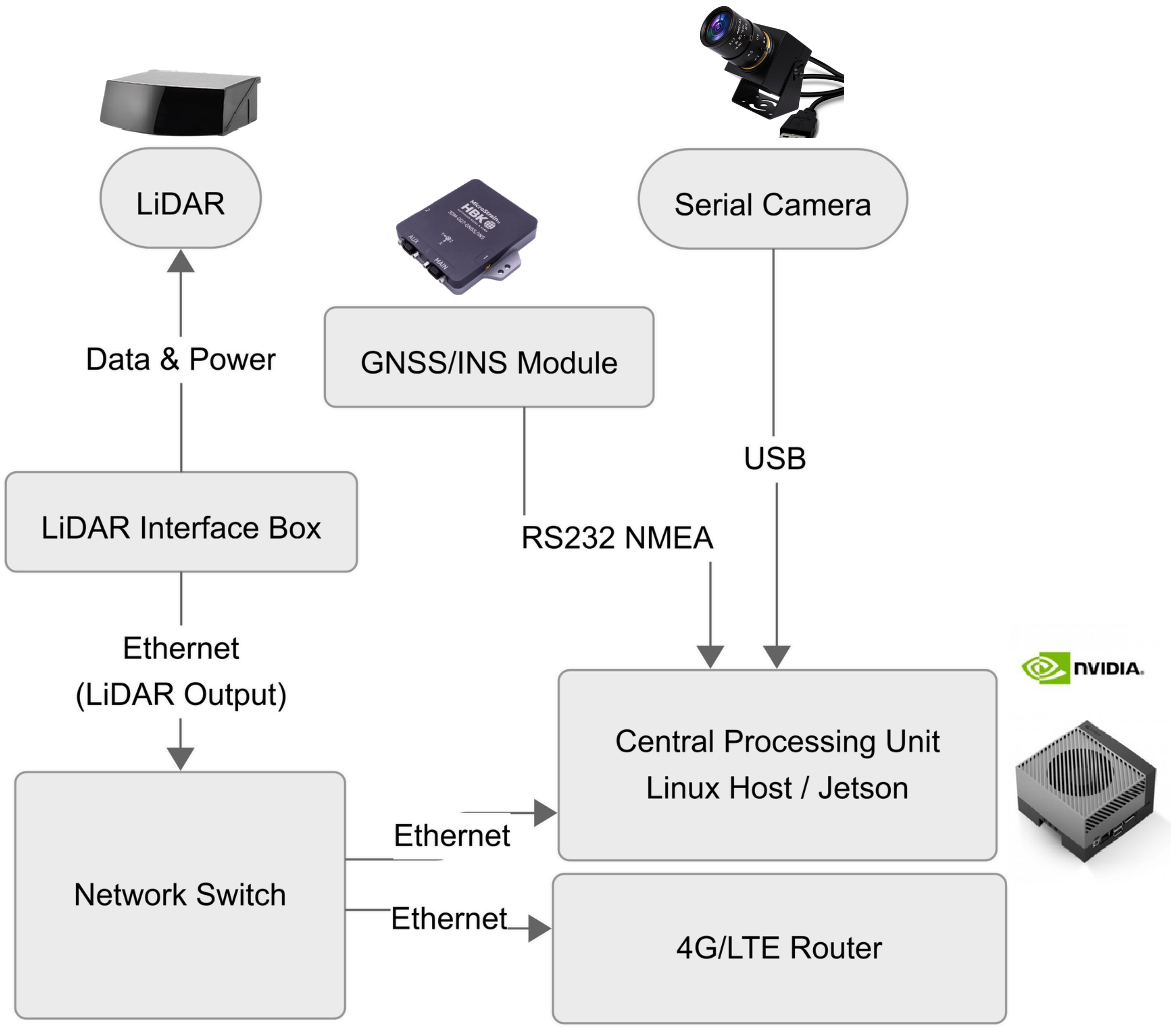
Orientation & Position

Multi-Sensor Fusion Technology

Outputs:

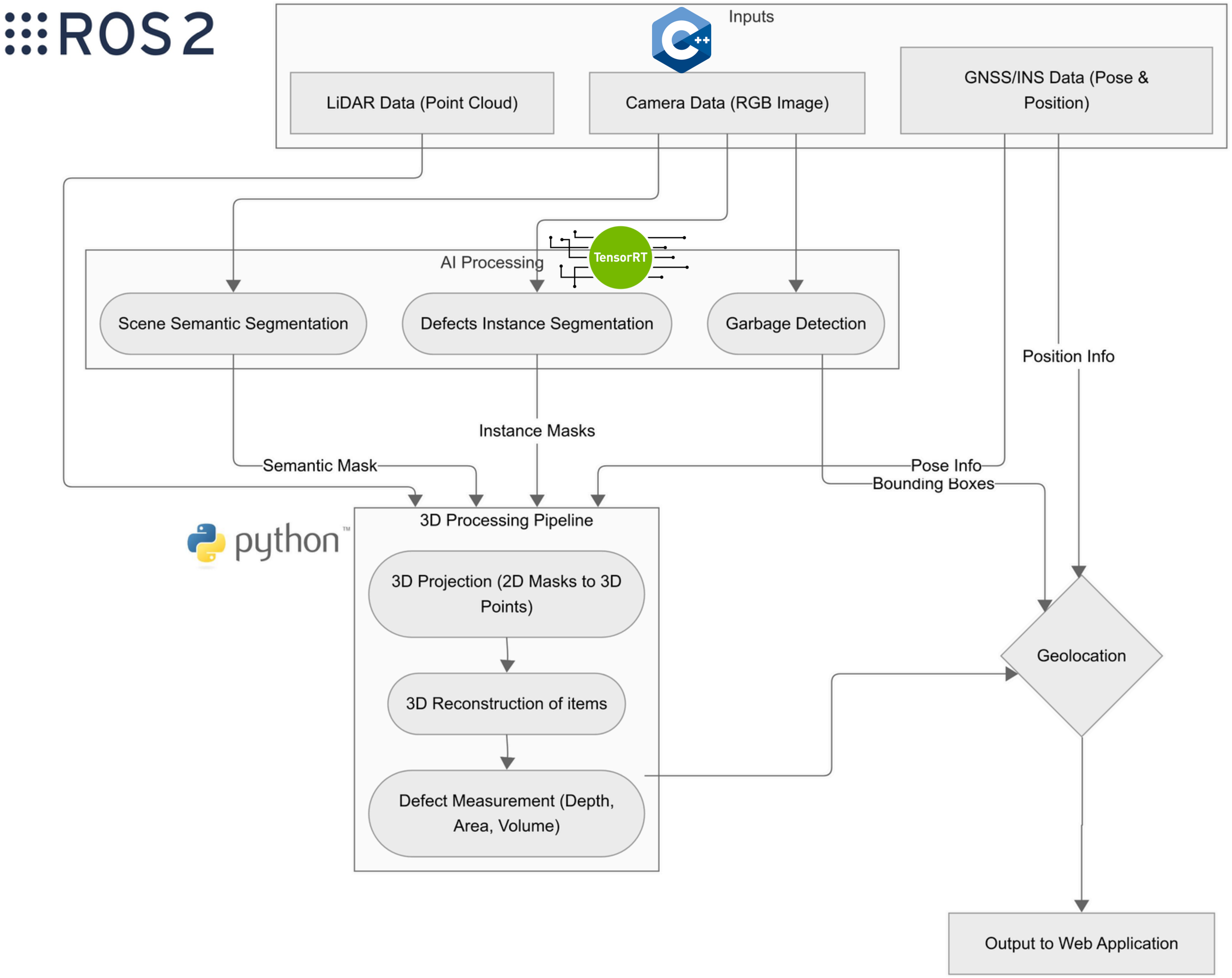
- **Detection of both pothole and cracks defects**
- **Scene understanding / relative positions**
- **3D reconstruction of detected items**
- **Estimation of Extension / Depth and Volume of detected items**
- **Accurate in-road positioning of each detected item**

Hardware Setup



A vehicle-mounted system designed for continuous and automated road monitoring

Software Architecture



Framework: Robot Operating System (ROS)

Nodes:

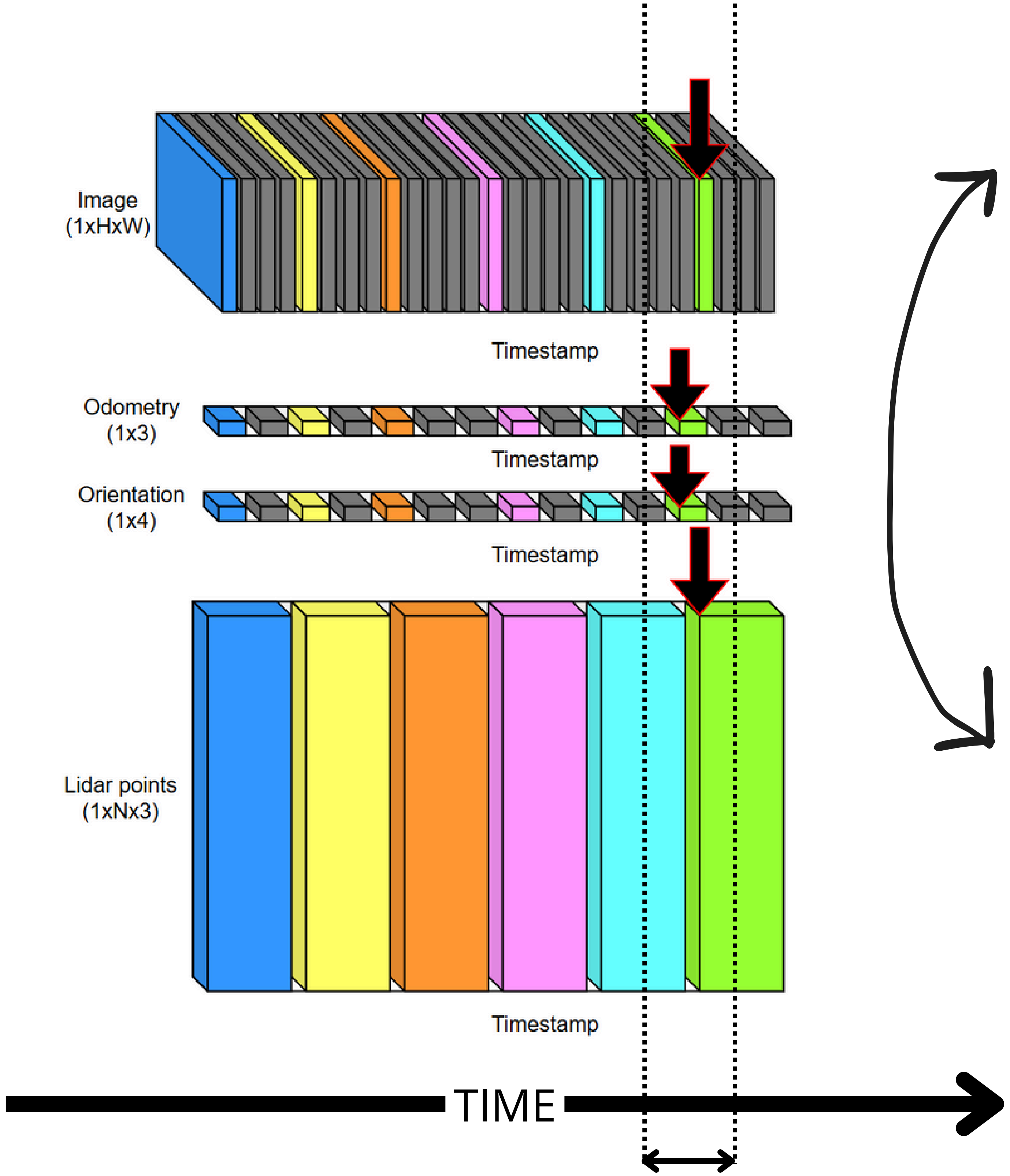
- Input **Sensors drivers Nodes** (Lidar, USB_Camera, GNSS/INS) → publish
- **Synchronization algorithm**
- ← subscribe, **AI Node**, gets RGB image and applies models:
  - Road defects instance segmentation
  - Scene semantic segmantation
  - Garbage detection → publish
- ← subscribe, **3D processing Node**, applies:
  - Camera - Lidar fusion
  - Point Cloud detection projection
  - Detected items reconstruction and measurement → publish
- ← subscribe, **Geolocation** and **data sending Node** → Cloud

Multi-sensor Synchronization

Challenge: *LiDAR, Camera, and GNSS/INS sensors operate at **different frequencies** and generate data **asynchronously**.*

Why it's Crucial: *Precise temporal alignment (synchronization) is essential for:*

- *Accurately projecting 2D detections (from Camera) onto the corresponding 3D point cloud (from LiDAR).*
- *Associating the correct vehicle pose (from GNSS/INS) with each sensor measurement.*
- *Ensuring reliable geolocation of detected road defects.*



tolerance: es. 100 ms

Bigger Tolerance:

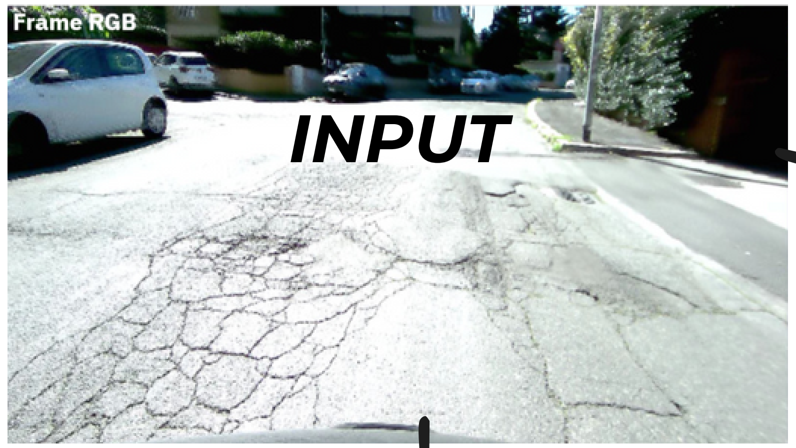
- *Risk to have **less sync***
- ***More** available data*

Tradeoff

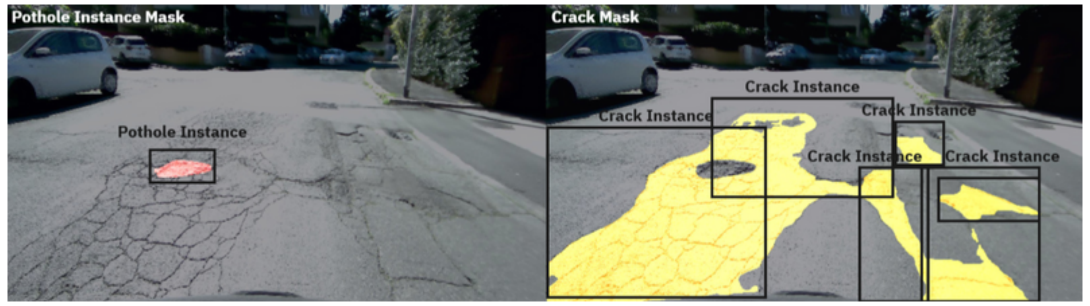
Lower Tolerance:

- *More **precise***
- *Risk to have **less data***

AI models applied to RGB frames

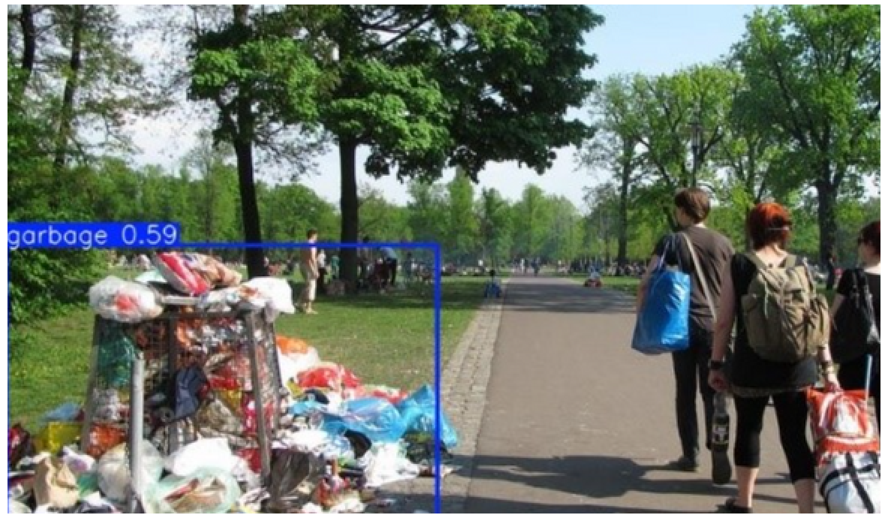


YOLO-Small Defects Instance Segmentation Architecture



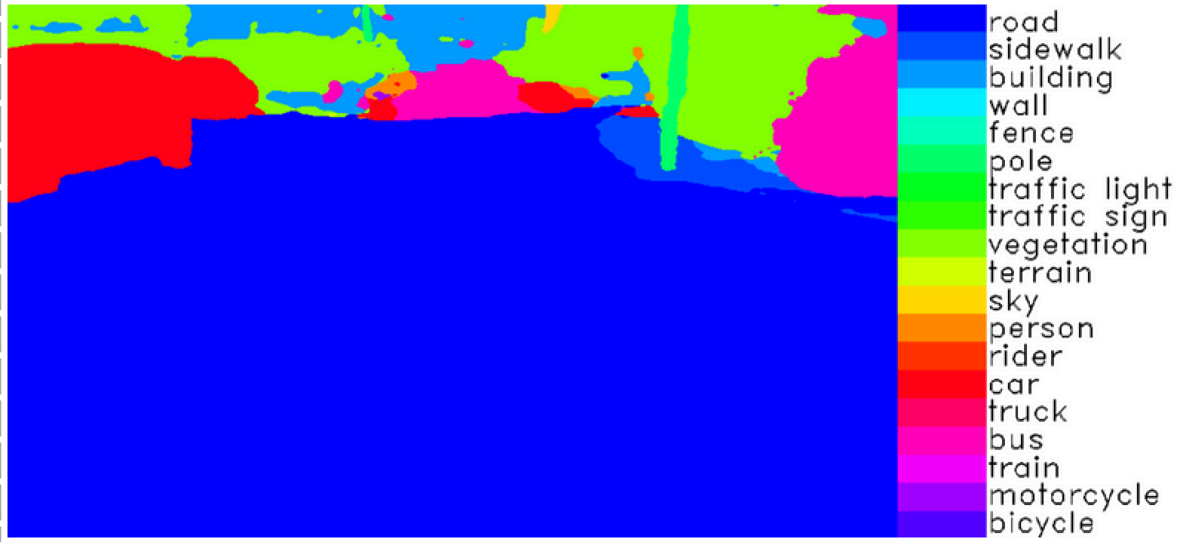
Output:
- *Masks and Bounding Boxes of each detected pothole (red)*
- *Mask of all cracks (yellow)*

YOLO-Nano Garbage Detection Architecture



Output:
- *Bounding Boxes of each detected garbage instance (blue)*

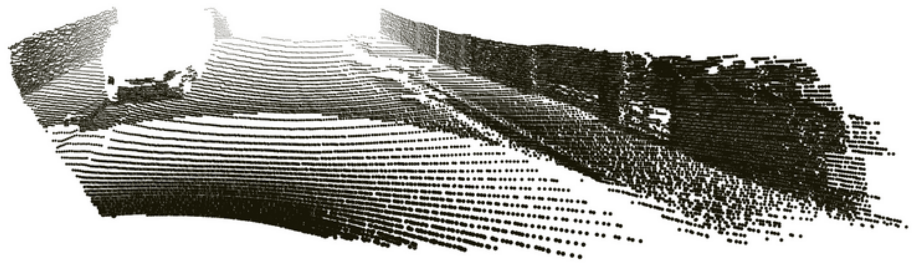
SegFormerB1 Scene Semantic Segmentation Architecture



Output:
- *Mask of semantic segmentation (each pixel is classified into specific classes)*

3D Lidar-Camera Fusion

Goal: To combine the strengths of both sensors, leveraging the AI's ability to detect defects in rich 2D camera images and the LiDAR's ability to provide precise 3D geometric measurements.



Requires:

- **Extrinsic Parameters**: The precise 3D position and orientation of the camera relative to the LiDAR sensor.

$$[R \mid t] = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \end{bmatrix}$$

PointCloud

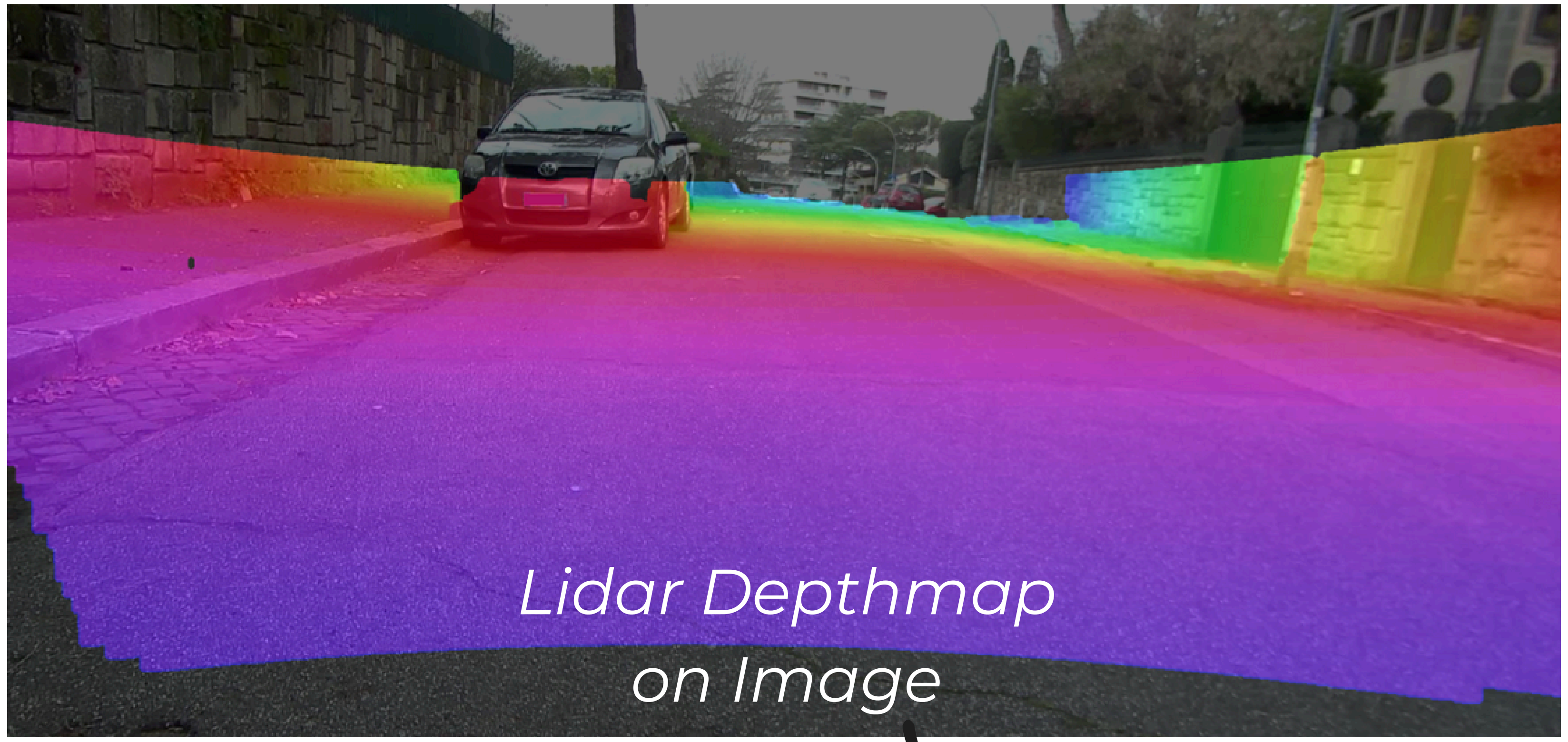
*CamWorld(x,y,z) = **RT***LidarWorld(x,y,z)*

- **Intrinsic Parameters:** Camera's internal characteristics (focal length, principal point, distortion).

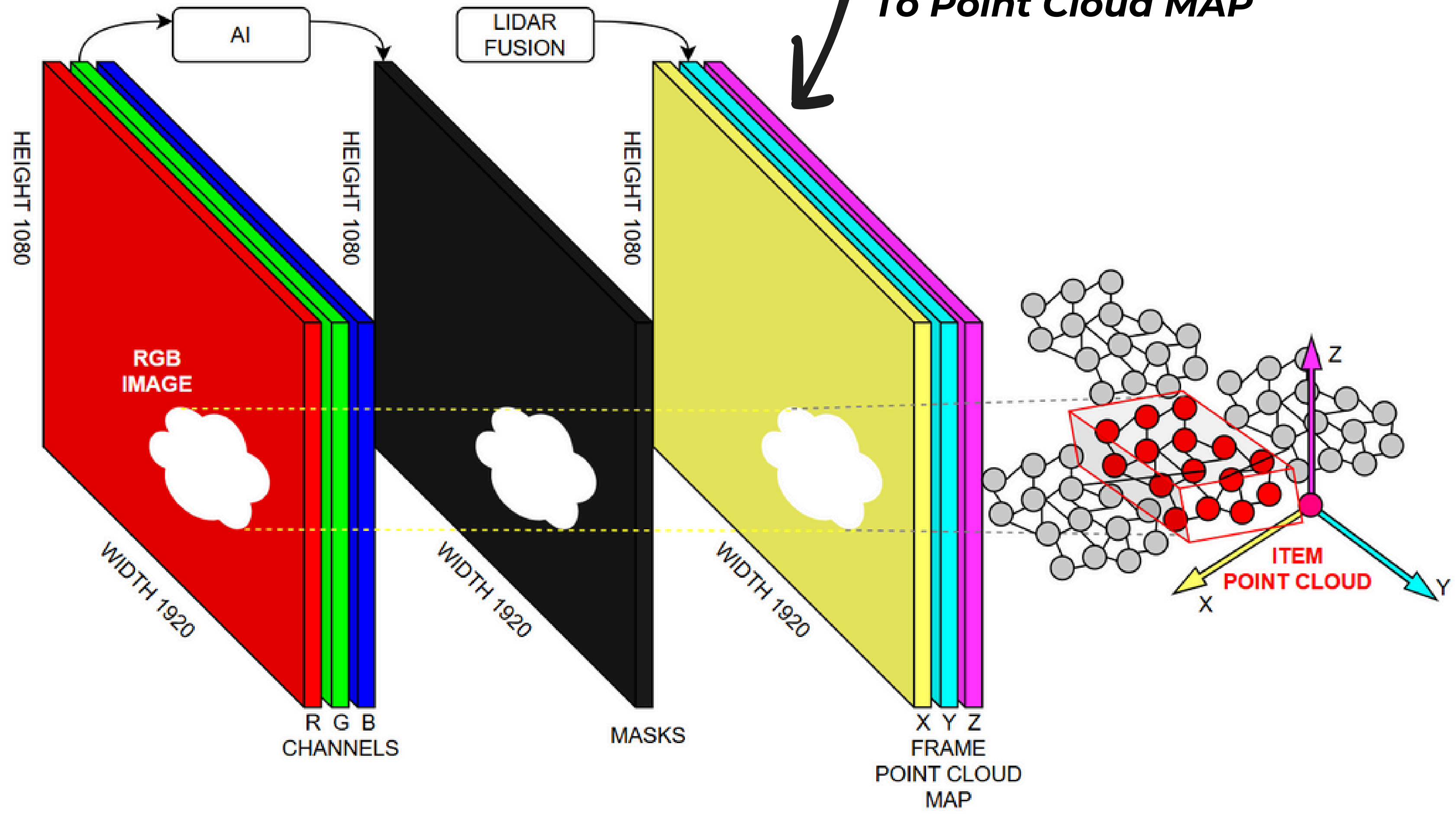
$$K = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

*Image(u,v) = **K***CamWorld(x,y,z)*

Depthmap



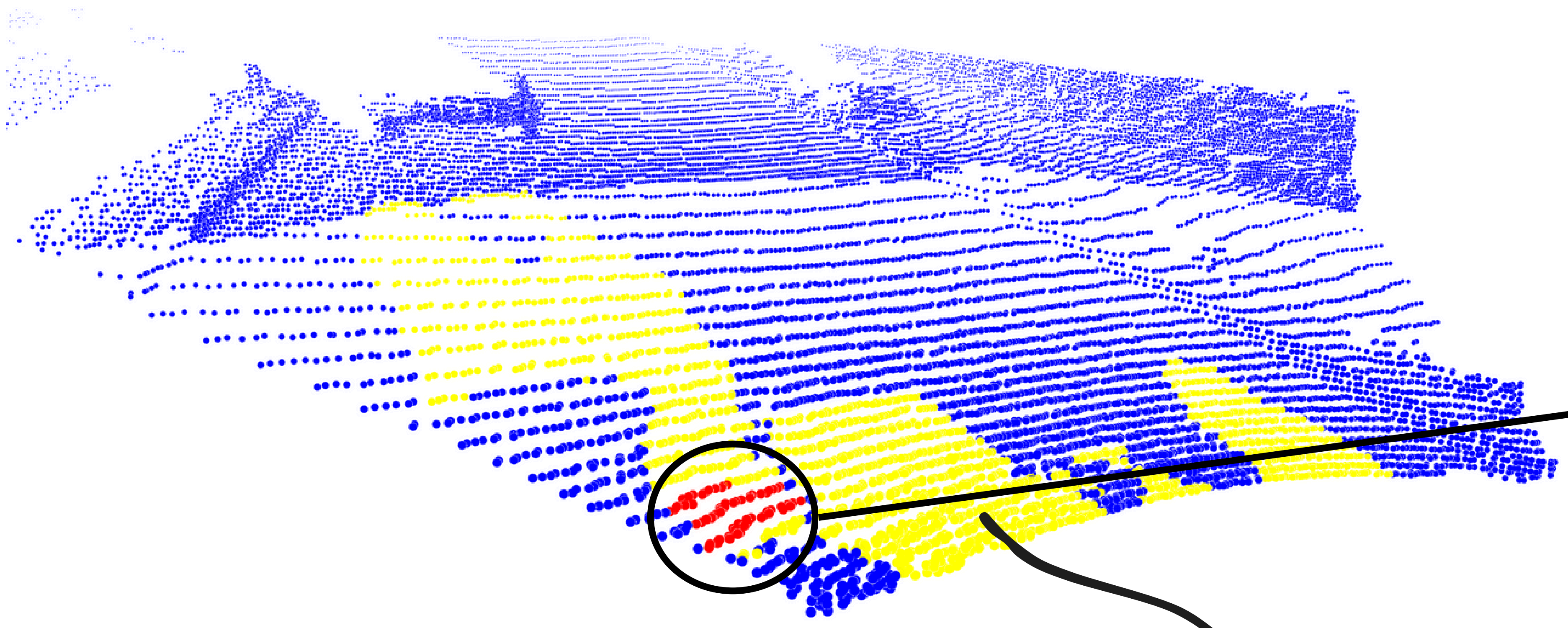
To Point Cloud MAP

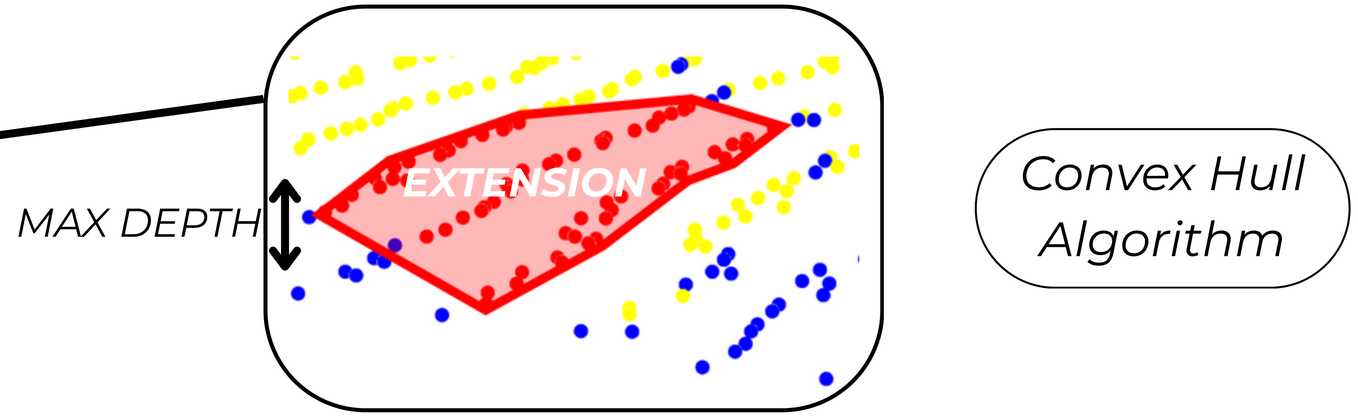


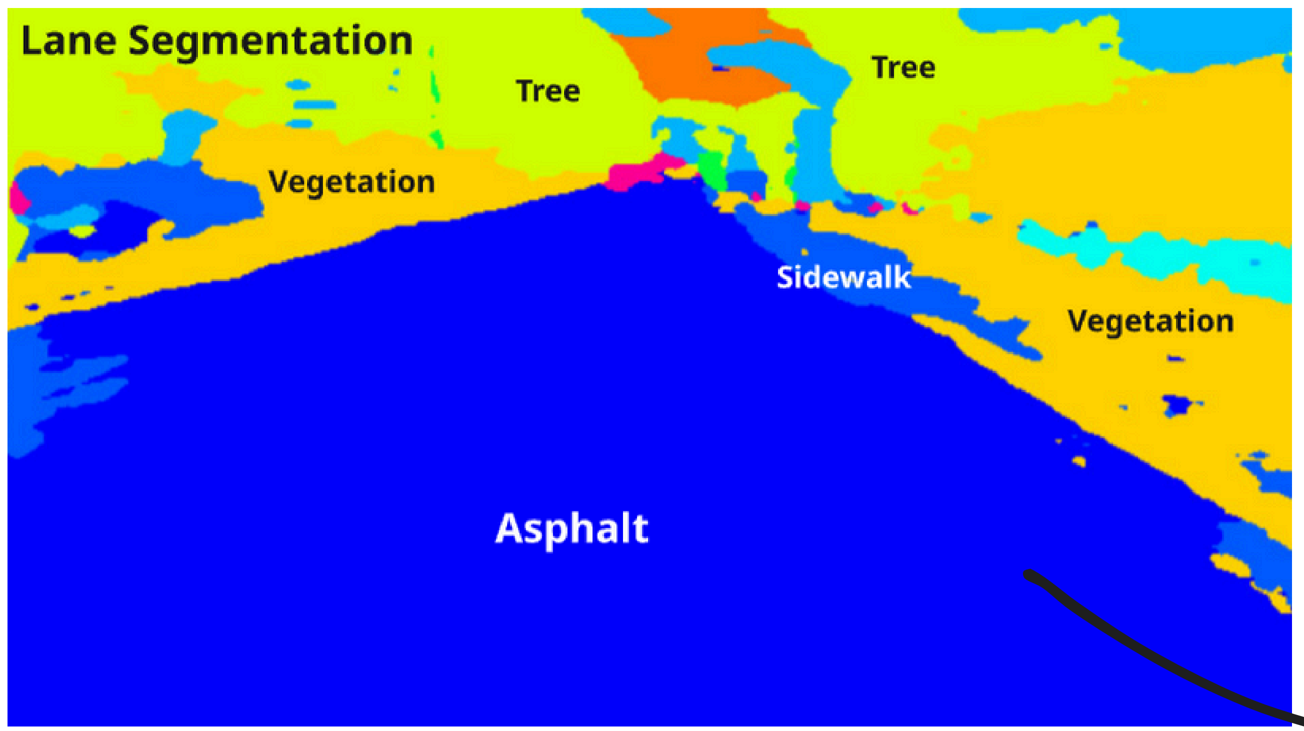
3D Measurements

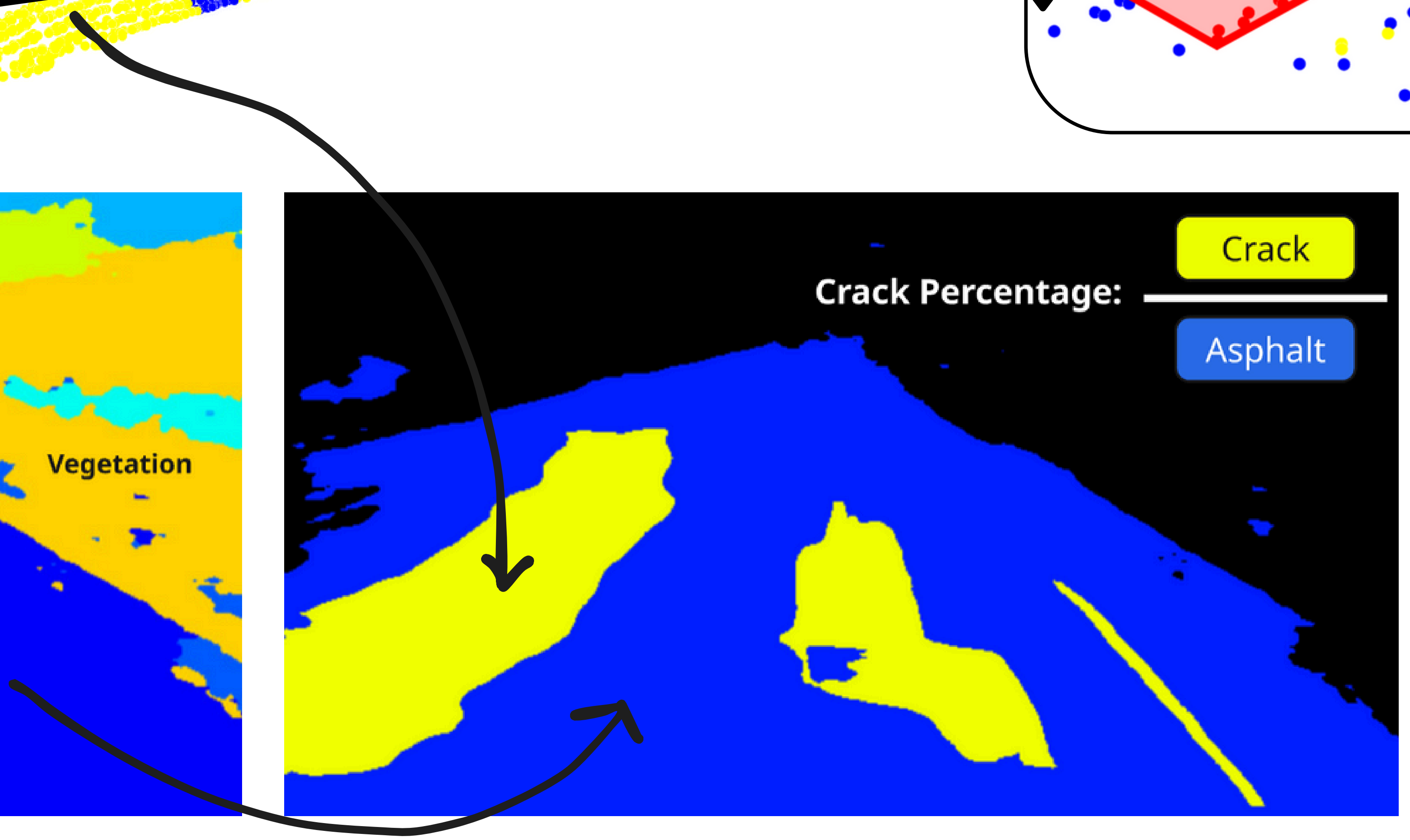
Once defects are projected onto Point Cloud domain, we can make some **measurements**:

- Pothole extension [cm2]
- Pothole max depth [cm]
- Pothole volume (how much asphalt) [cm3]









Crack mask and scene semantic segmentation are used together to compute **cracks percentage** over total asphalt:

Results: Defects segmentation AI performance

Validation Context*: The AI model was rigorously evaluated on dedicated test datasets, reflecting diverse urban road conditions.*

Performance metrics:

- **Mean Average Precision** (**mAP** @ IoU 0.5:0.95): **0.56**
  Interpretation: Reflects robust performance in detecting and segmenting potholes and alligator cracks across various sizes and appearances.

$$mAP = \frac{1}{k}\sum_{i}^{k} AP_i$$

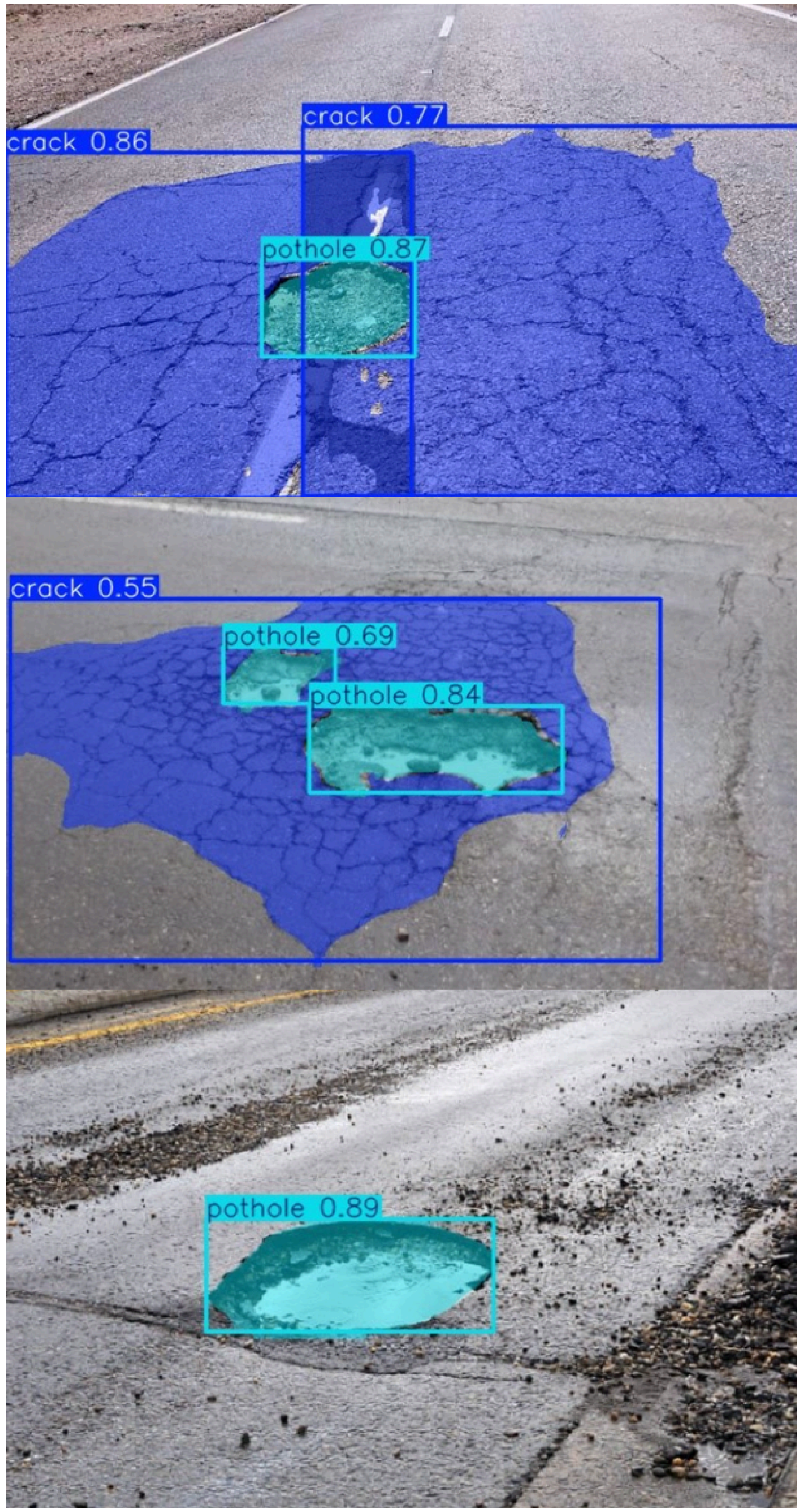
- ***F1-Score: 0.57***
  Interpretation: Indicates a solid balance between Precision (minimizing false detections) and Recall (capturing most actual defects).

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

TP = True positive
TN = True negative
FP = False positive
FN = False negative



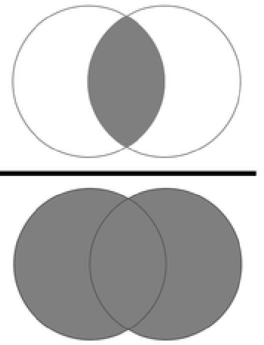
Results: Road segmentation AI performance

Validation Context: *The AI model was rigorously evaluated on dedicated test datasets, reflecting diverse urban road conditions.*

Performance metrics:

- **Mean Intersection over Union (mIoU): 0.43**
  Interpretation: Shows the model's capability to accurately outline the general road area within the scene.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



- **F1-Score (for 'Road' class): 0.98**
  Interpretation: Highlights exceptional reliability and precision in identifying the road surface itself, crucial for filtering out non-road detections.

	PREDICTED ROAD PIXELS	PREDICTED BACKGROUND PIXELS
TRUE ROAD PIXELS	0.99	0.01
TRUE BACKGROUND PIXELS	0.03	0.97

Results: 3D measurements performance

Validation Context: the performance of depth, extension and volume estimation has been evaluated by manual measuring (laser distance meter, on-site) a selection of real potholes.

3D Defect Dimensions:
Enabled by projecting AI detections onto the LiDAR point cloud for reconstruction.

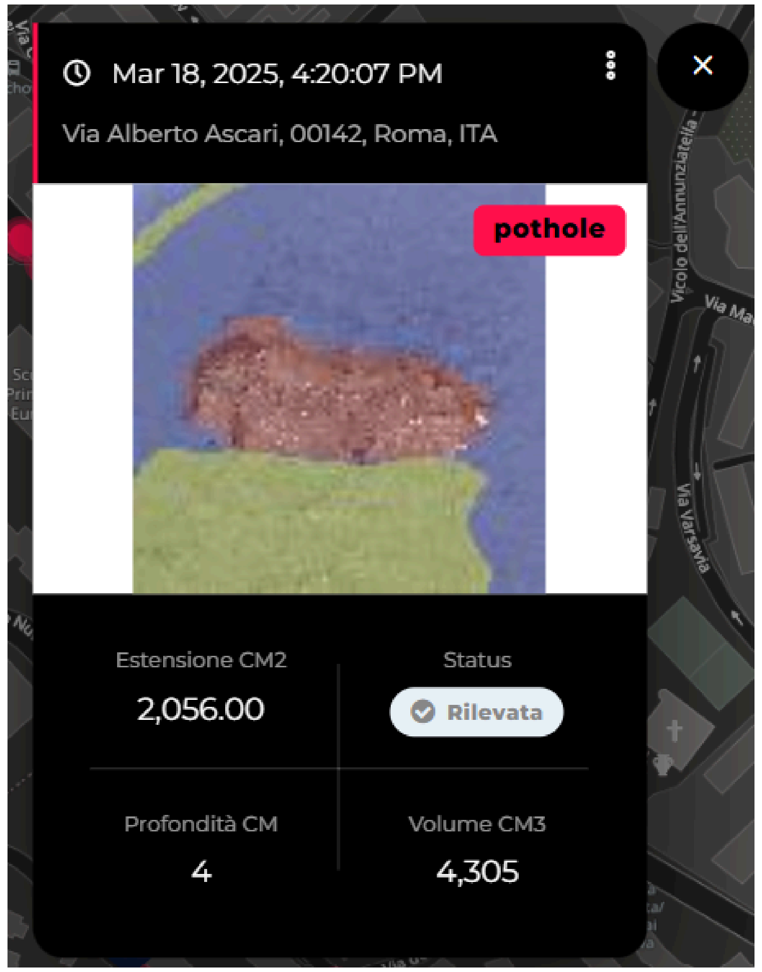
- **Depth:** relative error percentage of **24%**
- **Extension:** relative error percentage of **6%**
- **Volume:** relative error percentage of **19%**

Interpretation: Demonstrates quite high accuracy in quantifying the actual physical size of detected potholes. The extension is the most reliable measurement.

Possible sources of errors:

- **LiDAR intrinsic error**: +5cm/-5cm at long range
- **Sparse Point Cloud**: points decrease density at longer range
- **Motion effects**

Results: Web app navigation

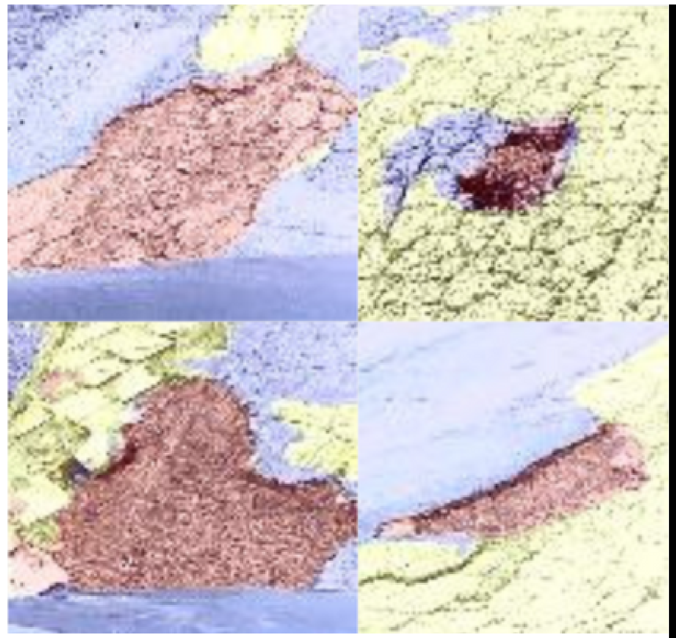


Information Box *for each detection:*

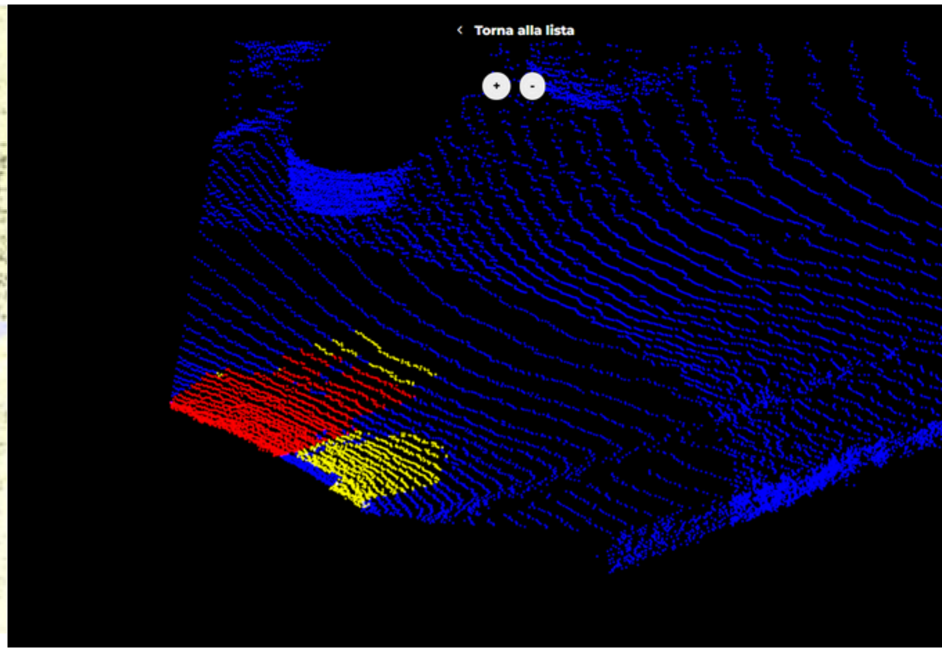
- *STATUS*
- *IMAGE*
- *EXTENSION*
- *DEPTH*
- *VOLUME*
- *COORDINATES*
- *DETECTION TIMESTAMP*

Heatmap of cracks: *which segments are the most affected*

Processed roads: *shows the areas that has been scanned with the system*

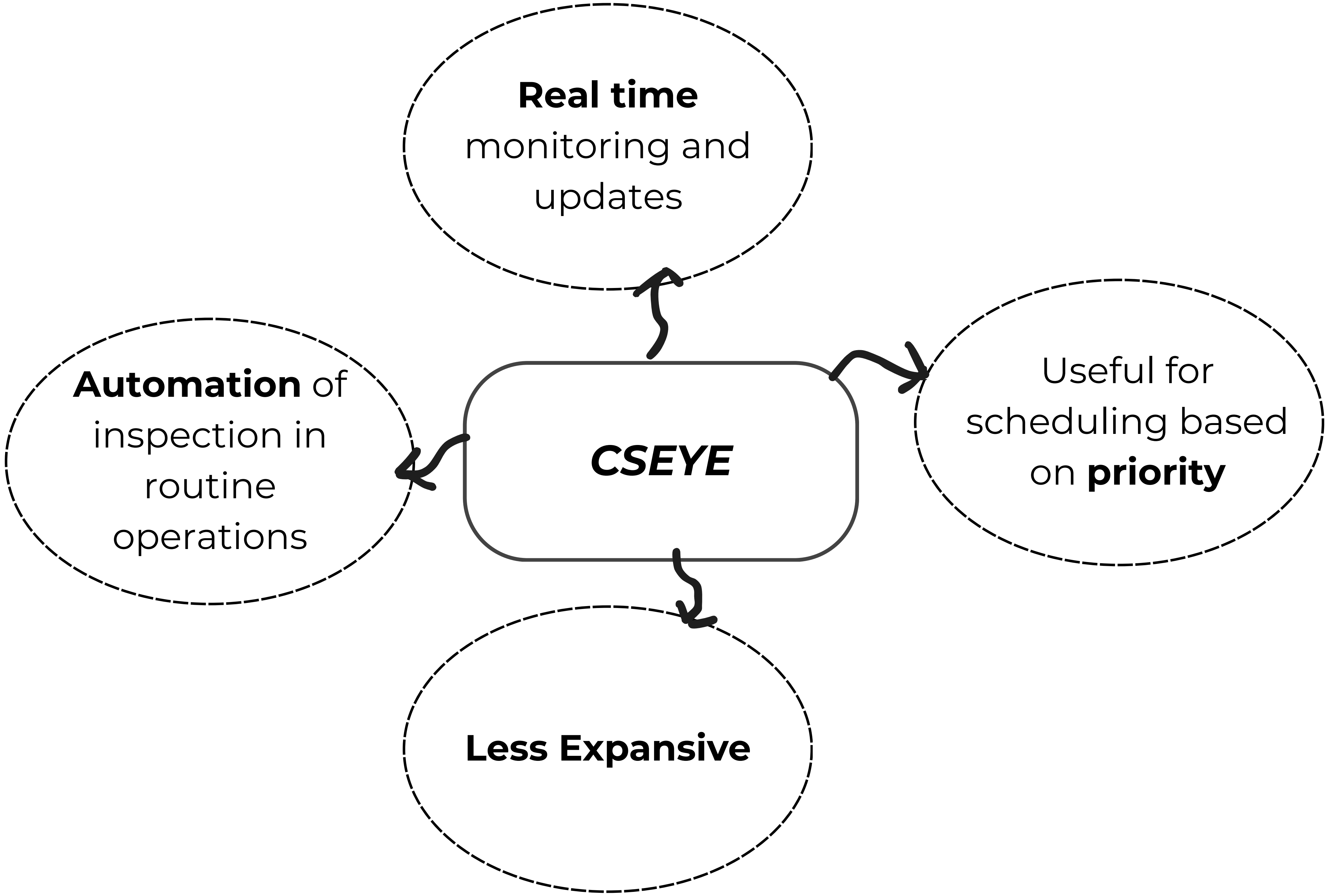


Images of each detection *(red mask-> pothole, yellow mask-> crack)*



3D Scene *point cloud with colorcoding*

Conclusions & Future Works



Future works:

- Sparse to Dense Depth Completion
- Exploring alternative approaches to LiDAR (mono and stereo depth estimation from Images)

KEY2

TOR VERGATA
UNIVERSITÀ DEGLI STUDI DI ROMA

Smart City Road Maintenance: A LiDAR and AI-Driven Approach for Detecting and Mapping Road Defects

Q&A
Thanks

Valencia, Spain April 06-10, 2025

IARIA