Uncertainty-Aware Road Obstacle Identification

Computer Vision / Project 11 - Prof. Amerini

Andrea Pizzi
Simone Turco
Federico Urbini

Outline:

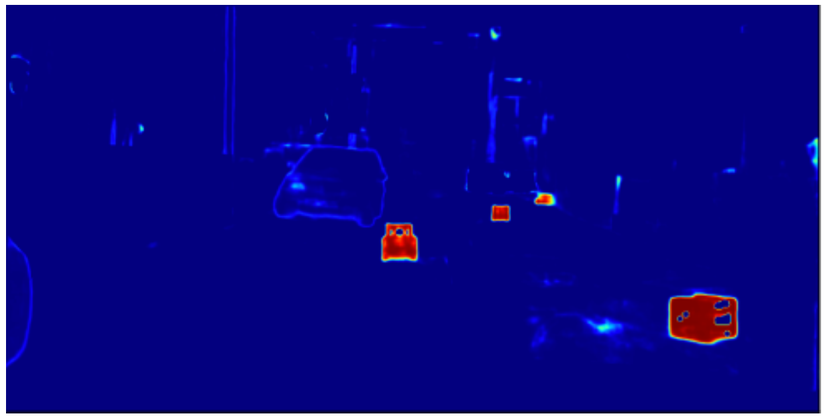
- Introduction
- Challenges
- Related Works
- Datasets
- Proposed Method
- Evaluation and Results
- Conclusions & Future Work

Why Uncertainty-Aware Road Obstacle Detection?

Traditional semantic segmentation models **assume** all objects belong to **known classes**.

Real-world driving involves **unexpected obstacles** (debris, boxes, unusual vehicles), so misclassifying unknowns as background leads to critical safety failures.

A model must understand what it doesn't know, **uncertainty** is key.



Why Add Uncertainty Awareness?

Uncertainty awareness means that a model not only makes **predictions** (e.g., where the road is) but also knows **how confident** it is in those **predictions**.

This **method** helps to:

- Improve **Safety**
- Highlights **low-confidence** regions
- Improve **Robustness**
- Adapt model to **Real Scenarios**

Challenges & Limitations in Semantic Segmentation

- **Small** or **distant** objects often lack sufficient pixel area to be detected reliably.

- **Domain shifts** like weather changes or unseen environments reduce generalization.

- **Unknown objects** may be misclassified as background.

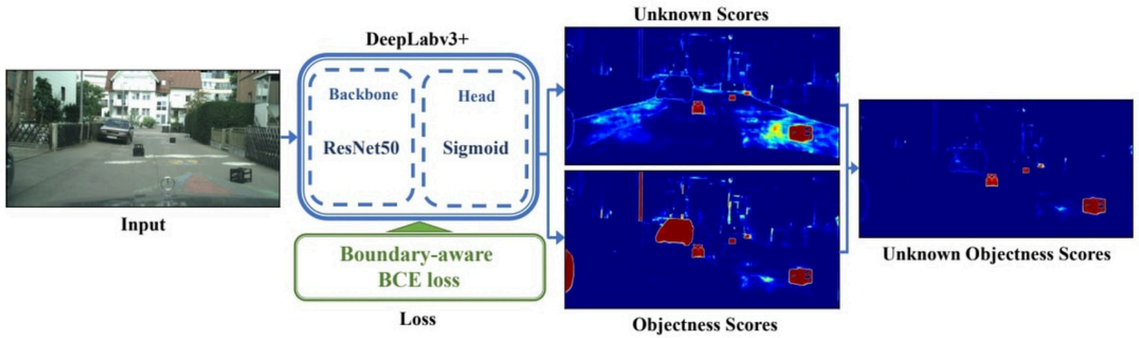
Related Works – Unknown Objectness and CP

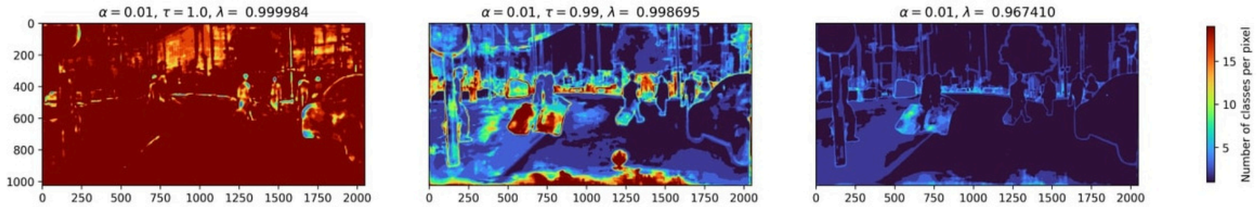
Noguchi et al. (2024) introduced UOS to improve OOD detection.

- Works independently of softmax thresholds.
- Highlights limitations of temperature scaling.

Mossina et al. (2024) proposed pixel-wise Conformal Prediction for semantic segmentation.

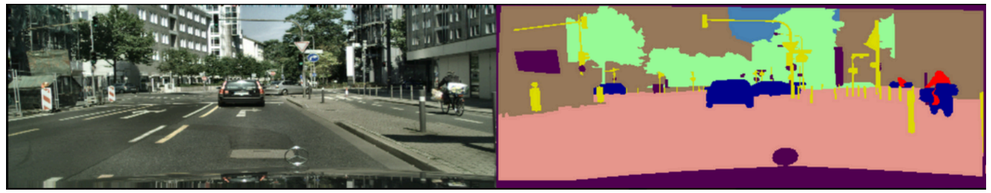
- Ensures a formal coverage guarantee (e.g., 95%) per pixel.
- Improves trust under distributional shift.





Datasets

Cityscapes



- Urban Scenes
- High-Quality/Corse Annotations.
- High Resolution (2048 × 1024 px)
- Large (more than 10.000)

LostAndFound

- Urban Scenes
- Very Small Obstacles
- High-Quality Annotations.
- High Resolution (2048 × 1024 px)

Proposed Method Overview

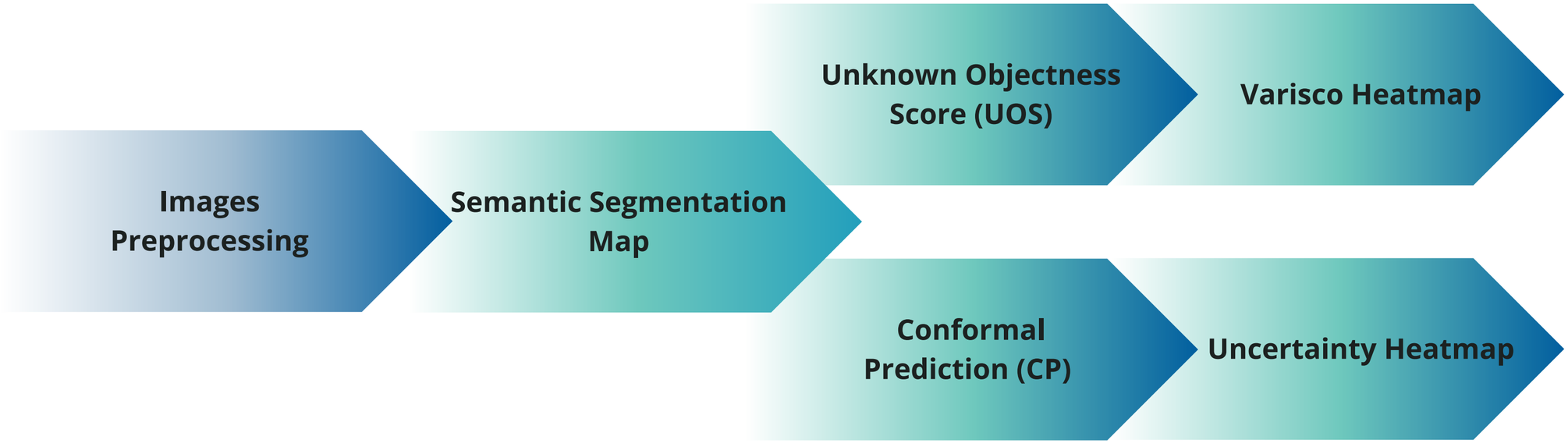


Image Preprocessing

- **Downscaling Dataset Images on a fixed size (512x256,)**
  Standardize input dimension and reduce memory usage and computational cost during traning
- **Contrast and brightness enhancement using CLAHE**
  CLAHE enhances visual clarity, especially in shadowed or overexposed regions
- **Implemented color-to-label mapping and reverse label-to-color mapping on the target image**
  Converts color-coded ground truth images into numerical class labels for training and enable reverse transformation to visualize prediction in colors
- **Transform intermediate mapping in one-hot encoding to match model logits**
  Aligns ground truth format with the model’s softmax/logit output

Semantic Segmentation and Object Class

We collapsed the original dataset classes into a reduced set of 8:

- simplify the segmentation task and help the model learn more effectively by focusing on fewer, more meaningful categories.

Introduced an 'object' class:

- Binary mask, where each pixel is assigned a value of 1 if it corresponds to an object of interest (e.g., vehicles, traffic signs, or road obstacles), and 0 otherwise.

Three steps for segmentation:

1. **Pre-training** on Extended Cityscapes with Coarse Annotations
2. **Training** on Cityscapes with Fine Annotations
3. **Fine-Tuning** on Lost and Found Dataset for Obstacle Detection

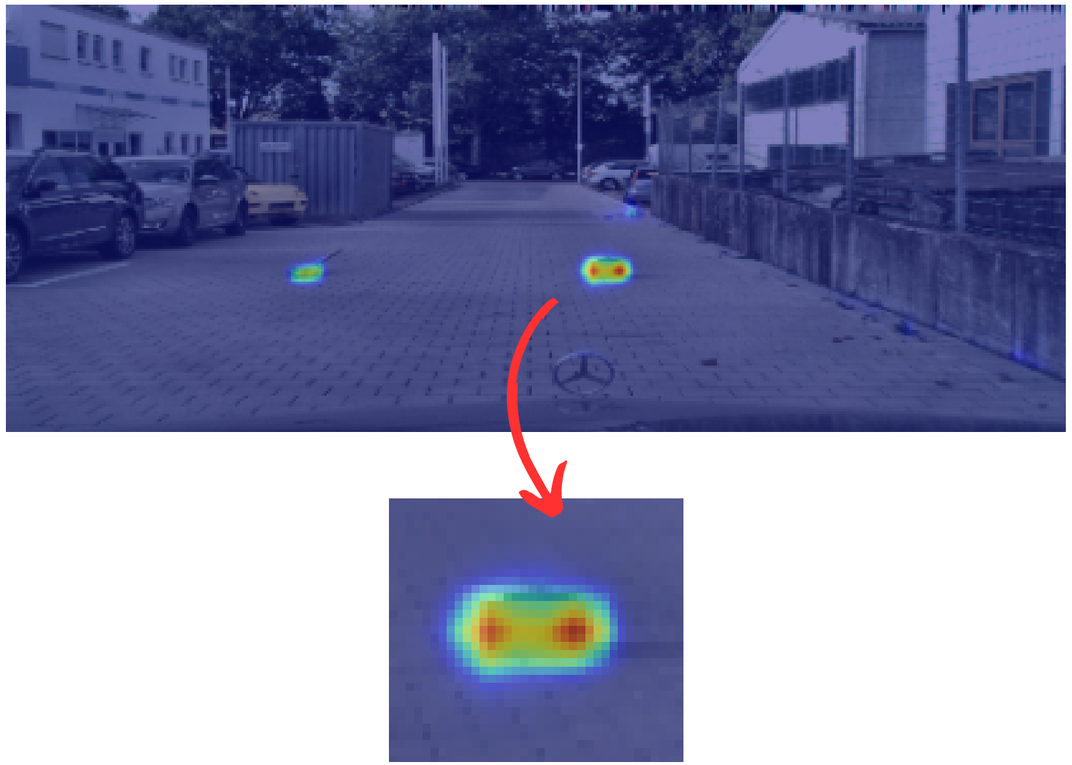
Unknown Objectness Score (UOS)

UOS is computed from softmax outputs as:

$$S_i = p_i^o \prod_{k=1}^{K} (1 - p_{ik}).$$

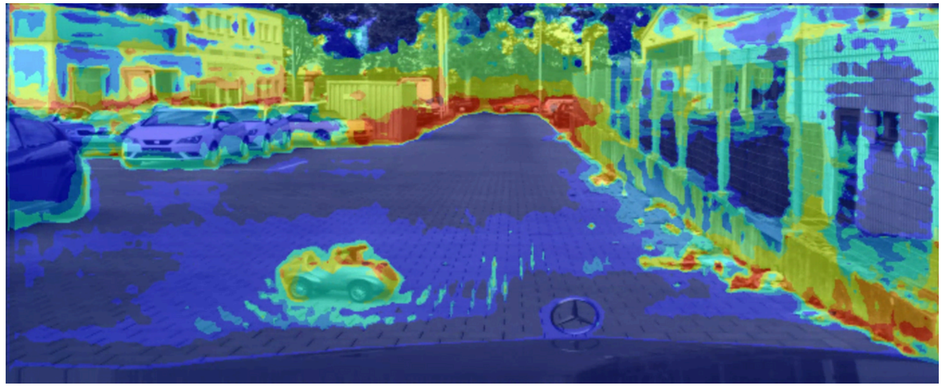
- Highlights pixels that do not belong confidently to any known class.
- Independent of softmax threshold → detects rare and out-of-distribution (OOD) objects.
- We used the Varisco heatmap to visualize Unknown Objectness Scores (UOS) as a color-coded overlay on RGB images.

Varisco Heatmap for Unknown Objectness Score



Conformal Prediction (CP)

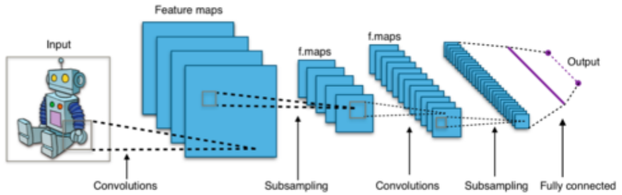
Conformal Prediction estimates how reliable a prediction is by comparing it to past examples. It helps identify uncertain regions where the model might be wrong.



lambda = 0.995

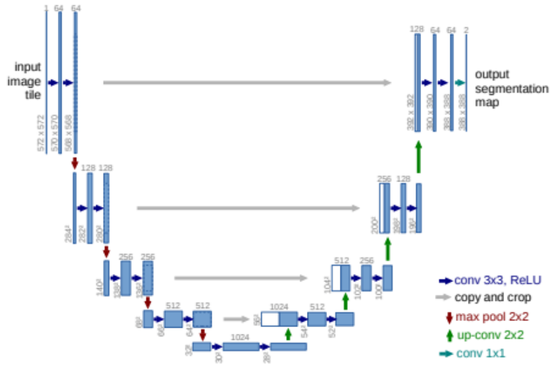
Proposed Networks

FCN Vgg-19 based



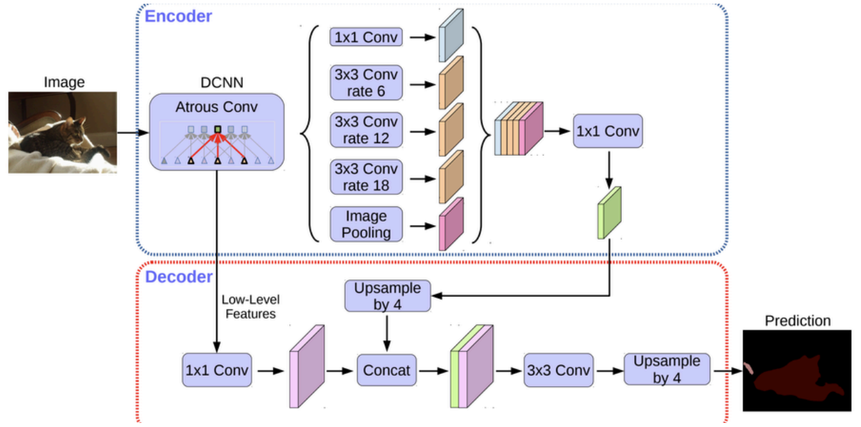
- **Full Convolutional** Network
- Standard Structure
- No Advance Features

U-Net



- **Encoder-Decoder** architecture
- **Weights share** at same level

DeepLabV3+



- **Multi-scale** segmentation
- Efficient dilatate convolution
- Network **State-Of-Art** for Segmentation Task

Evaluation Metrics, Benchmark

Train Metrics

- Segmentation: IoU, DiceScore

Benchmark metrics

- OOD Detection: AUROC, AP
- Uncertainty: Conformal coverage

Benchmark procedure

1. Test the model on LostAndFound Dataset
2. Repeate test 3 times
3. Estimate Uncertain
4. Collect Average results

Results

	optimizer	lr	Backbone	*Atrous rates*	*stride*	Decoder Depth	IoU	mAP	AUROC	Dice Score
DeepLab3+	SGD	0.01	ResNet50	[3,6,12]	16	5	0.554	0.2	0.93	0.62
U-Net	SGD	0.001	ResNet50	-	-	5	**0.57**	**0.46**	**0.96**	**0.65**
DeepLab3+	SGD	0.01	ResNet101	[3,6,12]	16	5	0.56	0.19	0.86	0.64

Among the tested architectures, U-Net consistently achieved the best performance on our task. proved highly effective, resulting in the highest scores across all evaluation metrics.

Conclusion & Future Work

Our integration of UOS and Conformal Prediction improves robustness under domain shift and enhances reliability. The framework is lightweight, modular, and generalizable.

Future Directions:

- Real-time deployment of UOS & CP
- Fusion with multi-modal data (RGB + depth, LiDAR) and network
- Use UOS maps to guide refinement: when the model is uncertain, activate a second-stage module to reprocess those regions and improve the final segmentation."

Contacts:

Andrea Pizzi, pizzi.1995517@studenti.uniroma1.it
Simone Turco, turco.1987712@studenti.uniroma1.it
Federico Urbini, urbini.2007465@studenti.uniroma1.it