

Figure 1: Label errors detected by our method in the Pascal VOC dataset.

Scope:

- The presence of label errors in datasets is well known
- Most learning methods aware of label errors aim at training well in presence of label errors
- Explicit detection of label errors has only been done in image classification [4]
- This work establishes a detection method for label errors in semantic segmentation datasets and a corresponding benchmark

Method:

- We use deep learning & uncertainty quantification (UQ) [2] to get proposals for missing labels / incorrect classes
- If a segment is a false positive (no match in the ground truth) and has high confidence (according to our UQ), this is a candidate to be reviewed; it might be a label error; see Fig. 2 for an illustration of our method

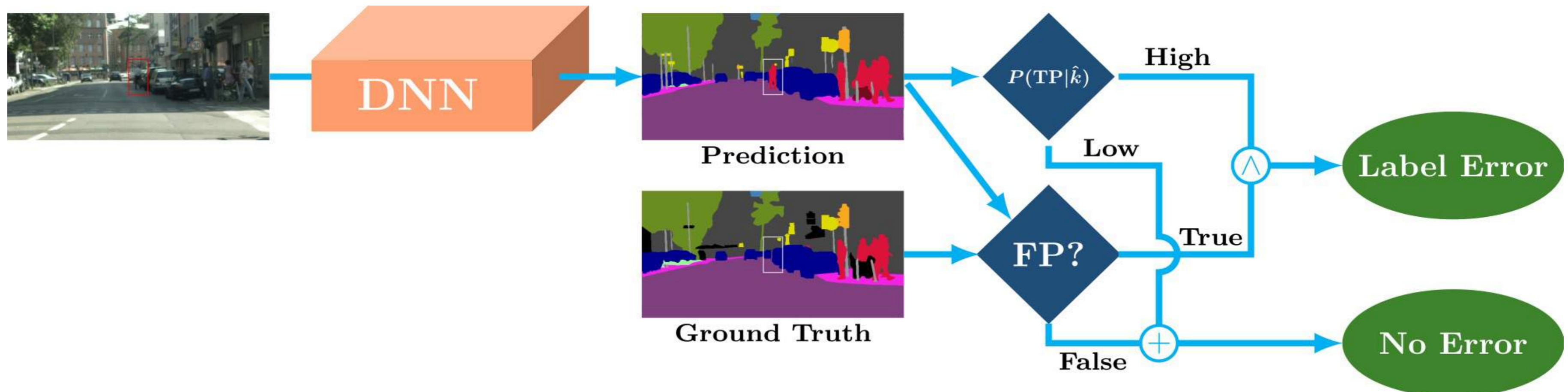


Figure 2: A sketch of our method. A false positive segment with high estimated confidence is sent to label error review.

© 2023 University of Wuppertal, please contact us before re-distribution

Benchmark and Results:

- We introduce a benchmark for evaluating label error detection by dropping segments from the ground truth and finding them with our method; for evaluation we utilize [3]
- The benchmark results show that we are consistently more efficient than when reviewing false positives randomly
- They also reveal that our method becomes more efficient compared to a random false positive review the less label errors are left in the dataset
- We reviewed 200 cases of highest confidence for different datasets in Table 1; see [1] for further results; see Fig. 1 for exemplary label errors found in Pascal VOC

References:

- [1] Rottmann, Matthias, and Marco Reese. "Automated Detection of Label Errors in Semantic Segmentation Datasets via Deep Learning and Uncertainty Quantification." WACV 2023.
- [2] Rottmann, Matthias, et al. "Prediction error meta classification in semantic segmentation: Detection via aggregated dispersion measures of softmax probabilities." IJCNN 2020.

Dataset	mIoU	TP	FP	Prec
Cityscapes Training	86.82	115	85	57.50
Cityscapes Validation	86.82	106	94	53.00
PascalVOC Validation	78.03	95	105	47.50
Coco-Stuff Validation	28.20	134	66	67.00
ADE20K	43.12	110	90	55.00

Table 1: Results for real label errors in datasets

Partners



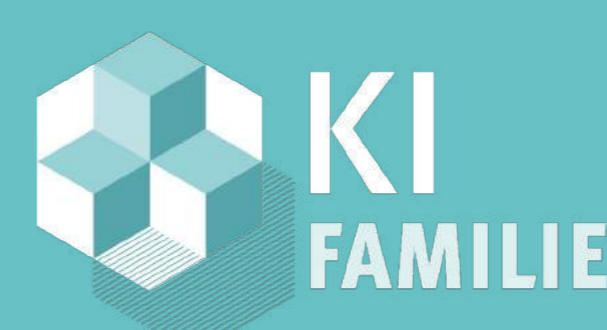
External partners



For more information contact:

rottmann@uni-wuppertal.de
marco.reese@de.bosch.com

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.



Supported by:



on the basis of a decision by the German Bundestag