

Active Learning for Infrastructure Distress Segmentation via Pixel-Level Difficulty Estimation

Hong Lang^{1,2}, Jiang Chen¹, Hongren Gong¹, Zhen Chen¹, Qingwen Zhou³, Jian John Lu^{1,4}

¹ Tongji University | ² University of Illinois Urbana-Champaign | ³ Wayne State University | ⁴ Shandong Jiaotong University

Contact Email: honglang@illinois.edu

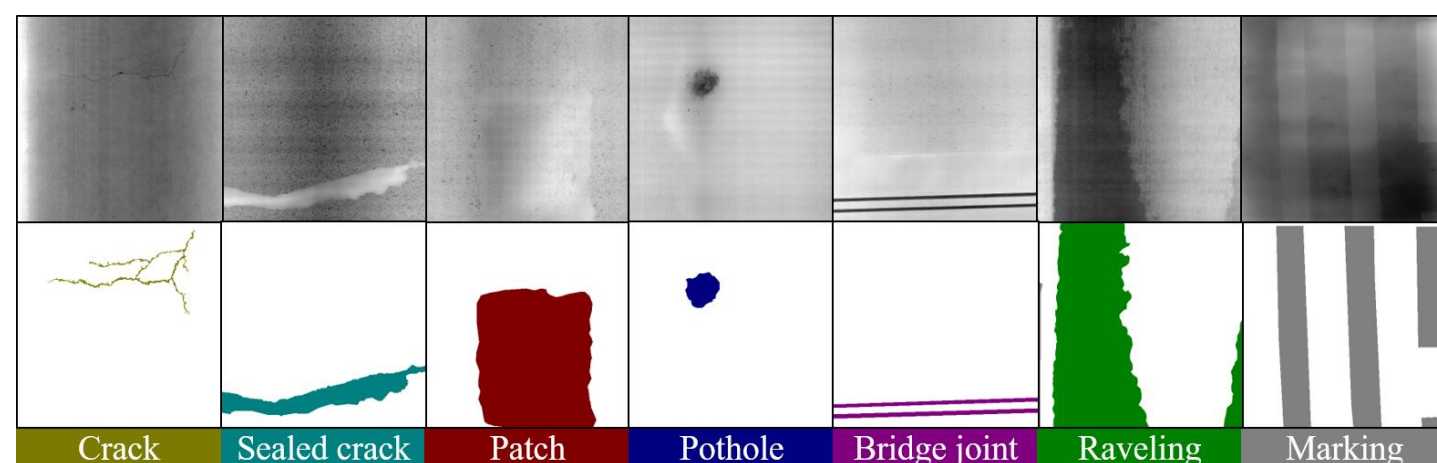


BACKGROUND & MOTIVATION

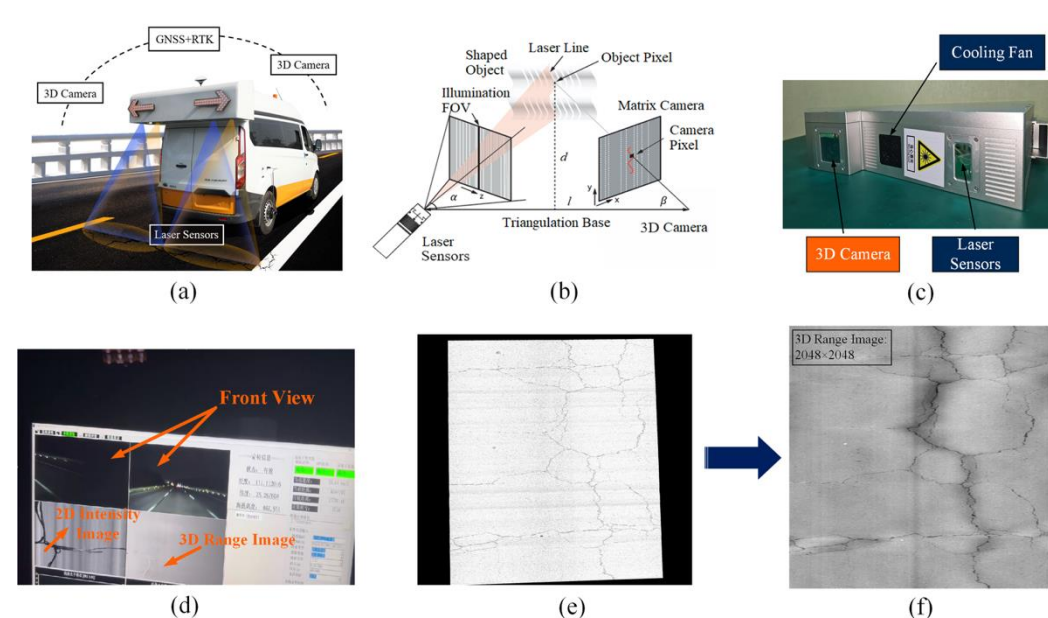
- Pavement distress threatens road safety and shortens service life.**
 - 3D pavement imaging adopted by many US states — captures sub-millimeter surface detail
 - Deep learning enables pixel-level distress segmentation across diverse road conditions
 - Pixel-level annotation is costly: 30–60 min per crack image for human experts
 - Low labeled-data regime severely limits model performance in real deployments

KEY QUESTION: Can we achieve high segmentation accuracy using only a fraction of labeled data?

PAVEMENT DATASET & DATA COLLECTION



Representative high-precision 3D surface data collected by 3D laser system



- 3,000 3D Pavement Images;**
- 8 Semantic Categories;**
- 5,000+ km Surveyed Length**

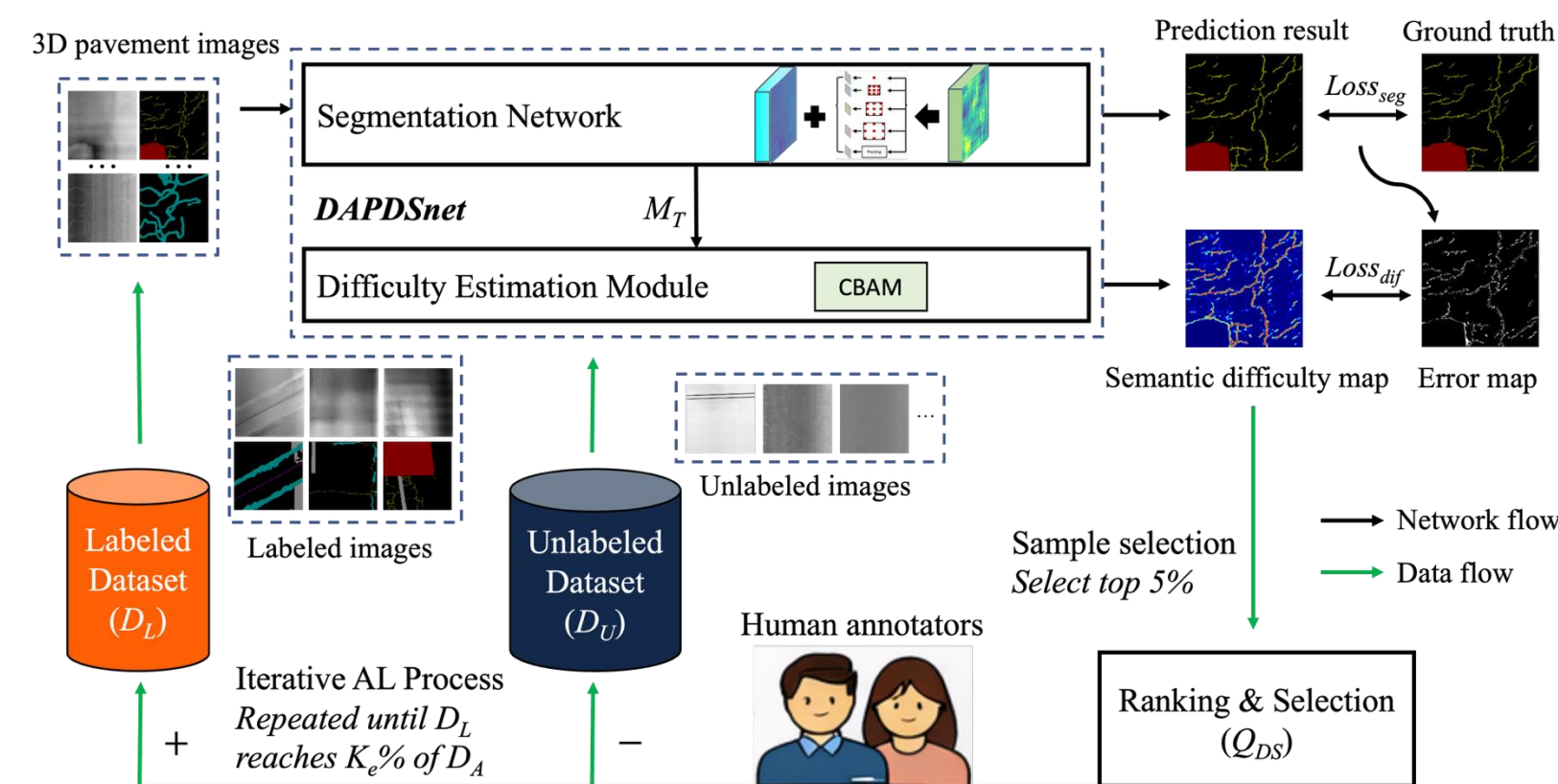
ACTIVE LEARNING (AL) PROCESS

- Initialize:** Small random labeled pool D_L (5–20%)
- Train:** Train DAPDSnet on current D_L
- Infer:** Run model on unlabeled pool D_U → difficulty maps
- Rank & Select:** Score Q_{DS} = avg pixel difficulty → pick top 5%
- Annotate:** Human experts label selected samples → D_L grows

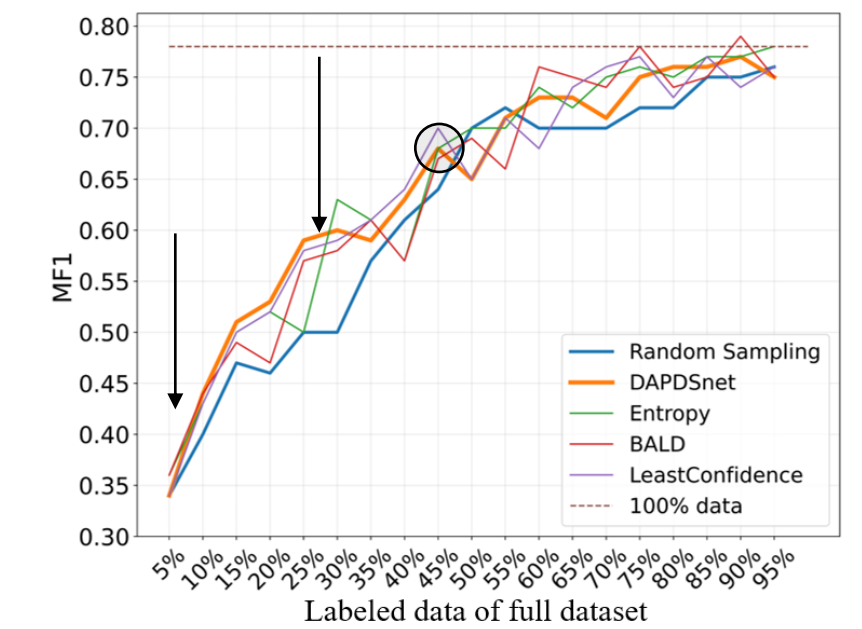
Repeat until D_L reaches target coverage $K_e\%$ of full dataset D_A

DAPDSnet FRAMEWORK & AL Baseline

Overall Active Learning Framework



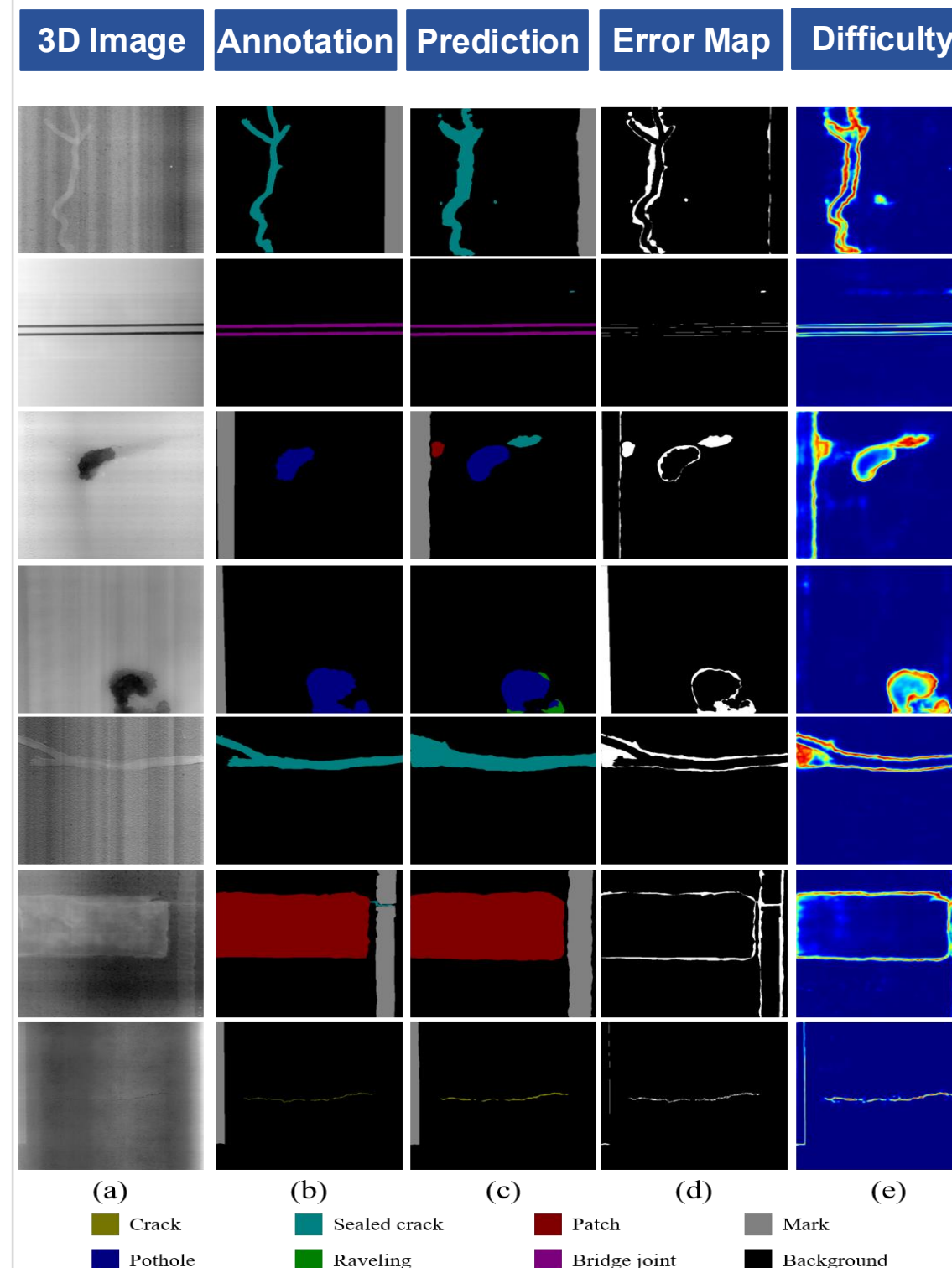
At 30% labeled data: DAPDSnet 0.6 MF1 vs. Random 0.5 MF1 (+20% relative gain)



89% of Full MF1 at 45% data #1 vs. 4AL Baseline (0~45%)

QUALITATIVE RESULTS

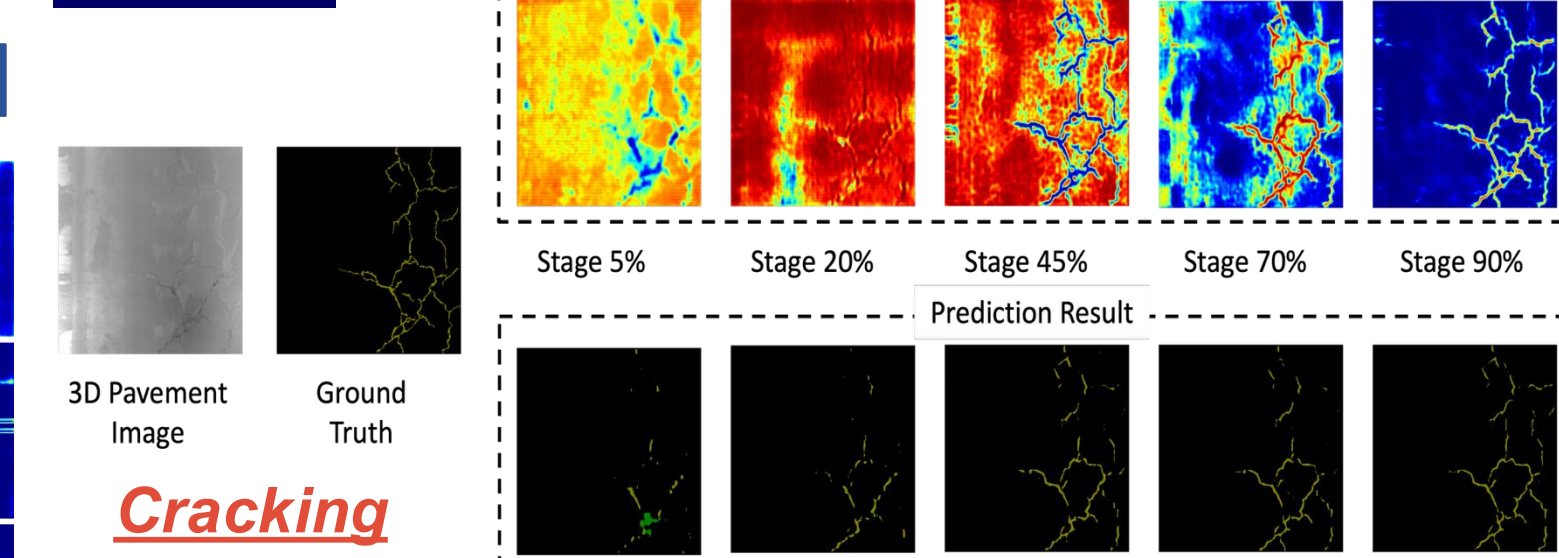
Prediction at ~90% Labeled Data Stage:



Segmentation aligns closely with annotations across all distress types.

AL PROCESS VISUALIZATION

How difficulty maps evolve as more data is labeled:



SUMMARY

- DAPDSnet achieves **~85% MIoU** & **~89% MF1** of fully-supervised performance using only **45%** labeled data
- MobileNetV2 backbone: computationally efficient (**17ms/image**) and well-suited for iterative AL deployment
- Consistently outperforms Random Sampling, and AL baselines in the critical low-data regime (**5%–40%**)
- Broadly applicable to **infrastructure inspection** tasks where pixel-level labeling is the primary bottleneck

Acknowledgements: NSFC, HENAN-DOT, HUALU TECH, PRES