arXiv - arXiv:2403.17010
Home - ldkong.com/Calib3D
GitHub - ldkong1205/Calib3D

Calib3D: Calibrating Model Preferences for Reliable 3D Scene Understanding

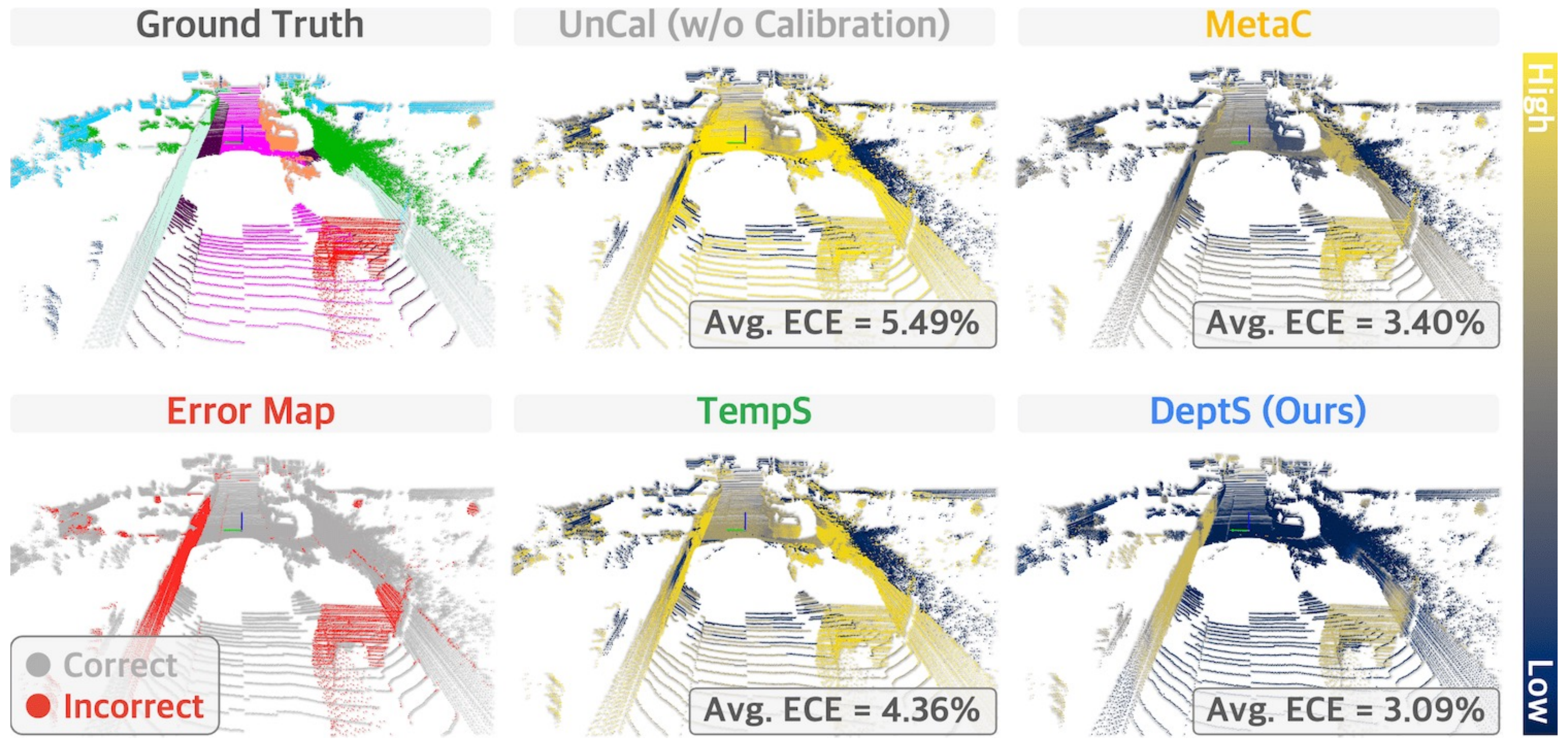
Lingdong Kong* Xiang Xu* Jun Cen Wenwei Zhang Liang Pan Kai Chen Ziwei Liu

WACV 2025
TUCSON, ARIZONA • FEB 28 - MAR 4

Motivation & Contribution

TL;DR

➢ **Calib3D** is a large benchmark for examining the **uncertainty** of 3D scene understanding systems under different driving conditions.



➢ Well-calibrated 3D scene understanding models are anticipated to deliver **low uncertainties** when predictions are **accurate** and **high uncertainties** when predictions are **inaccurate**.

➢ Our proposed depth-aware scaling (**DeptS**) is capable of outputting accurate estimates, highlighting its potential for **real-world usage**.

➢ The plots shown are the **point-wise calibration errors**. The color goes from dark to light denotes low and high errors, respectively.

NUS National University of Singapore
南京航空航天大學 NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS
上海人工智能实验室 Shanghai Artificial Intelligence Laboratory
香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

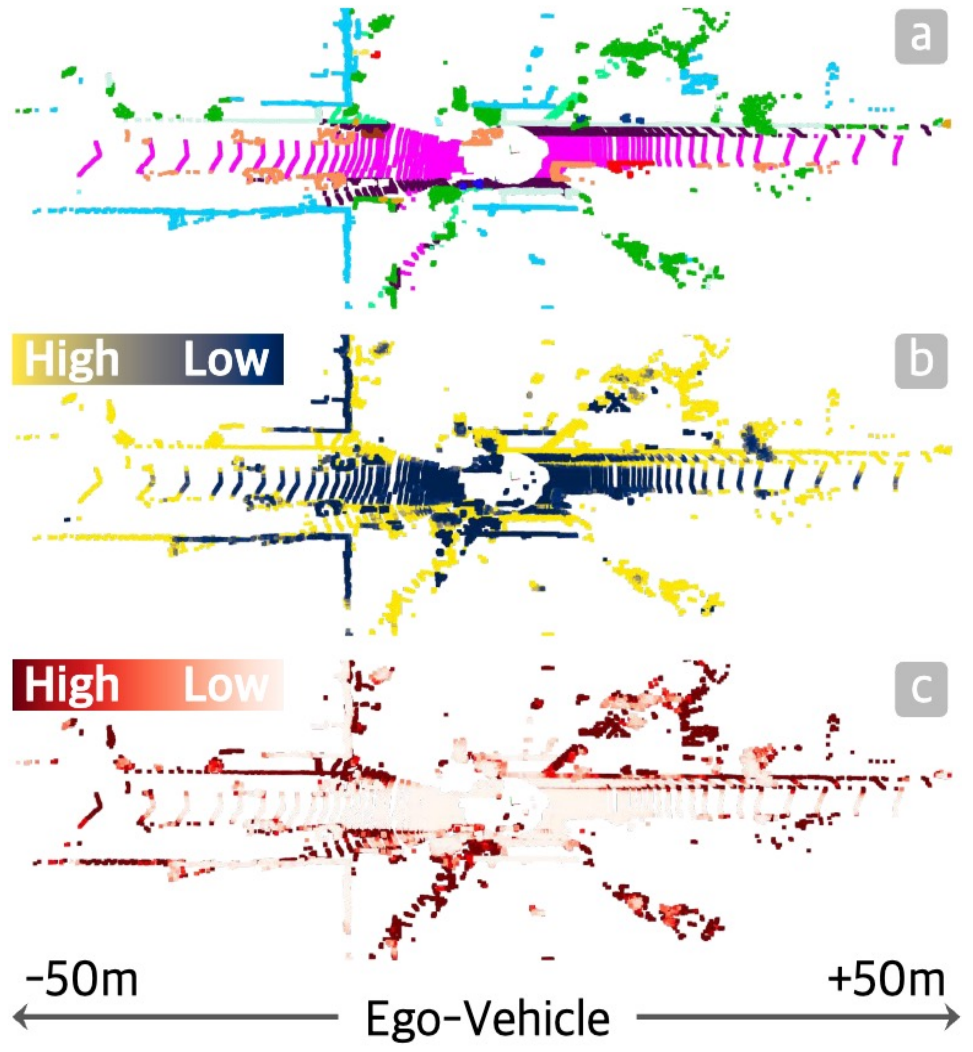
Design & Methodology

Benchmark Design

➢ 3D data are well-known to be highly **heterogeneous** due to different sensor acquisitions, placements, scene conditions, annotation protocols, etc.

➢ A learning system trained on different data sources tends to exhibit diverse **predictive confidence** and **accuracy**.

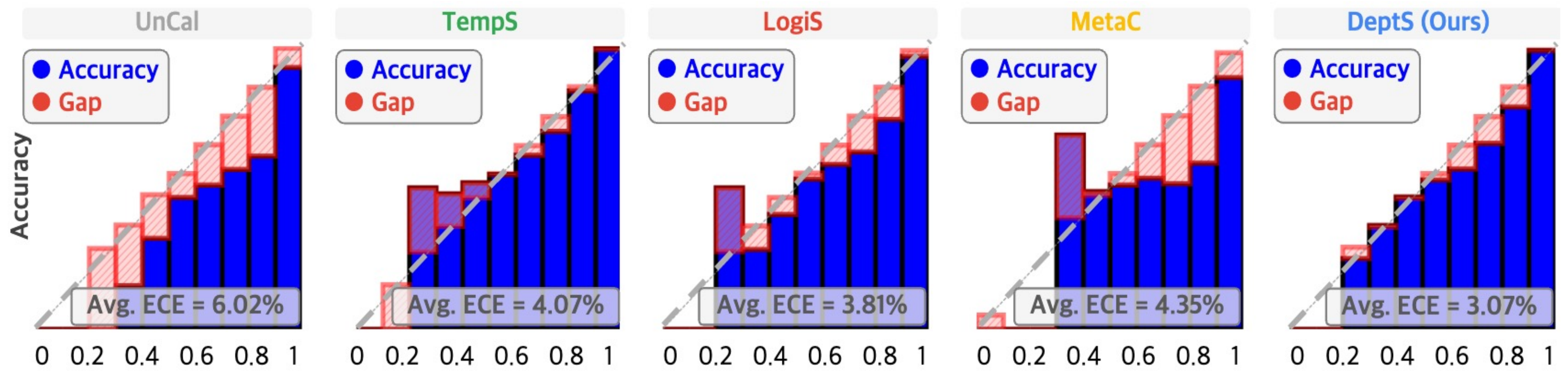
➢ **Calib3D** encompasses a study of 28 state-of-the-art models on 10 large-scale 3D datasets.



DeptS: Depth-Aware Scaling for Better 3D Calibration

➢ Different from RGB images, the point cloud data are **unordered** and **texture-less**, which inherits extra difficulties in feature learning.

➢ To pursue better 3D calibration, we propose a straightforward yet effective **DeptS** method. We observe a close correlation among calibration error, prediction entropy, and depth – an inherent 3D information that can be calculated based on Cartesian coordinates.



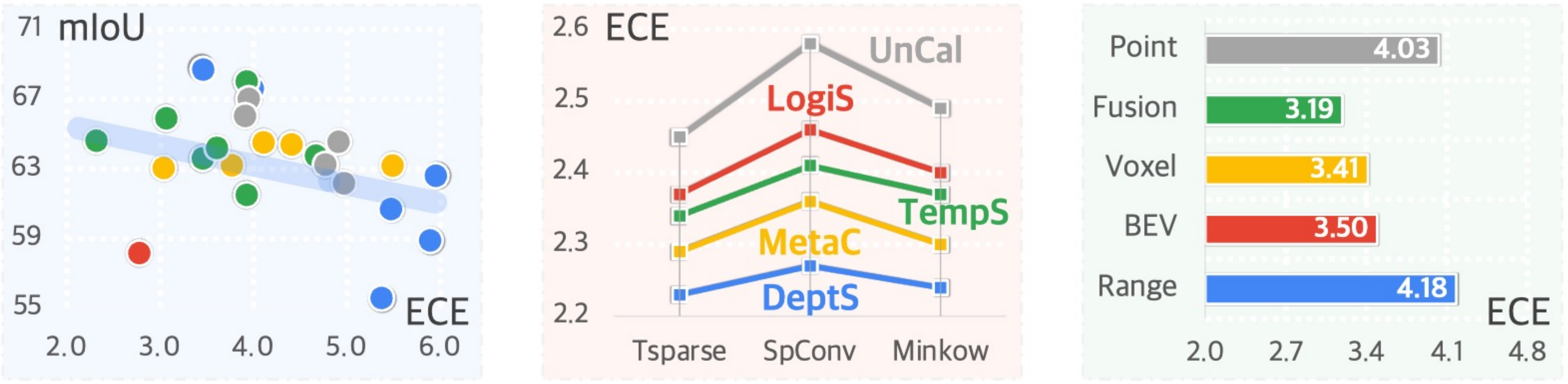
➢ Our approach is motivated by the observation that uncalibrated 3D models tend to have low accuracy in the **middle-to-far region** of the ego-vehicle, while posing severe **over-confident predictions**.

Experiments & Analysis

Comparative & Ablation Study

➢ **Aleatoric Uncertainty in 3D**. We examine how intrinsic factors, e.g., LiDAR sensor measurement noises and point cloud density variations, contribute to the data uncertainty in 3D models.

➢ **Epistemic Uncertainty in 3D**. Our investigation extends to model uncertainties coped with the diverse 3D architectures, highlighting the importance of addressing knowledge gaps in model training.



a. ECE vs. mIoU b. SparseConv Backend c. LiDAR Modality

➢ Through extensive evaluations of 3D models across diverse 3D datasets, we highlighted critical challenges in delivering **confident** and **accurate** predictions, particularly in safety-critical applications.

Type	nuScenes-C UnCal	nuScenes-C TempS	nuScenes-C LogiS	nuScenes-C DiriS	nuScenes-C MetaC	nuScenes-C DeptS	SemanticKITTI-C UnCal	SemanticKITTI-C TempS	SemanticKITTI-C LogiS	SemanticKITTI-C DiriS	SemanticKITTI-C MetaC	SemanticKITTI-C DeptS
Clean •	2.45%	2.34%	2.34%	2.42%	2.29%	2.23%	3.04%	3.01%	3.08%	3.30%	2.69%	2.63%
Fog ○	5.52%	5.42%	5.49%	5.43%	4.77%	4.72%	12.66%	12.55%	12.67%	12.48%	11.08%	10.94%
Wet Ground ○	2.63%	2.54%	2.54%	2.64%	2.55%	2.52%	3.55%	3.46%	3.54%	3.72%	3.33%	3.28%
Snow ○	13.79%	13.32%	13.53%	13.59%	11.37%	11.31%	7.10%	6.96%	6.95%	7.26%	5.99%	5.63%
Motion Blur ○	9.54%	9.29%	9.37%	9.01%	8.32%	8.29%	11.31%	11.16%	11.24%	12.13%	9.00%	8.97%
Beam Missing ○	2.58%	2.48%	2.49%	2.57%	2.53%	2.47%	2.87%	2.83%	2.84%	2.98%	2.83%	2.79%
Crosstalk ○	13.64%	13.00%	12.97%	13.44%	9.98%	9.73%	4.93%	4.83%	4.86%	4.81%	3.54%	3.48%
Incomplete Echo ○	2.44%	2.33%	2.33%	2.42%	2.32%	2.21%	3.21%	3.19%	3.25%	3.48%	2.84%	2.19%
Cross Sensor ○	4.25%	4.15%	4.20%	4.28%	4.06%	3.20%	3.15%	3.13%	3.18%	3.43%	3.17%	2.96%
Average •	6.78%	6.57%	6.62%	6.67%	5.74%	5.56%	6.10%	6.01%	6.07%	6.29%	5.22%	5.03%

➢ The results from **Calib3D** benchmark expose a significant gap in the calibration of current 3D models, which often achieve high **accuracy** but struggle to align **confidence** with predictive accuracy.