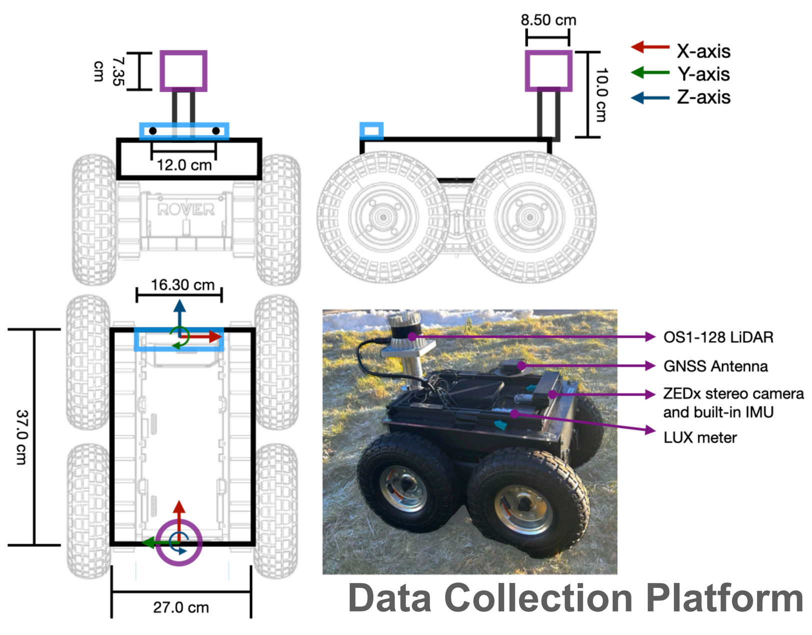
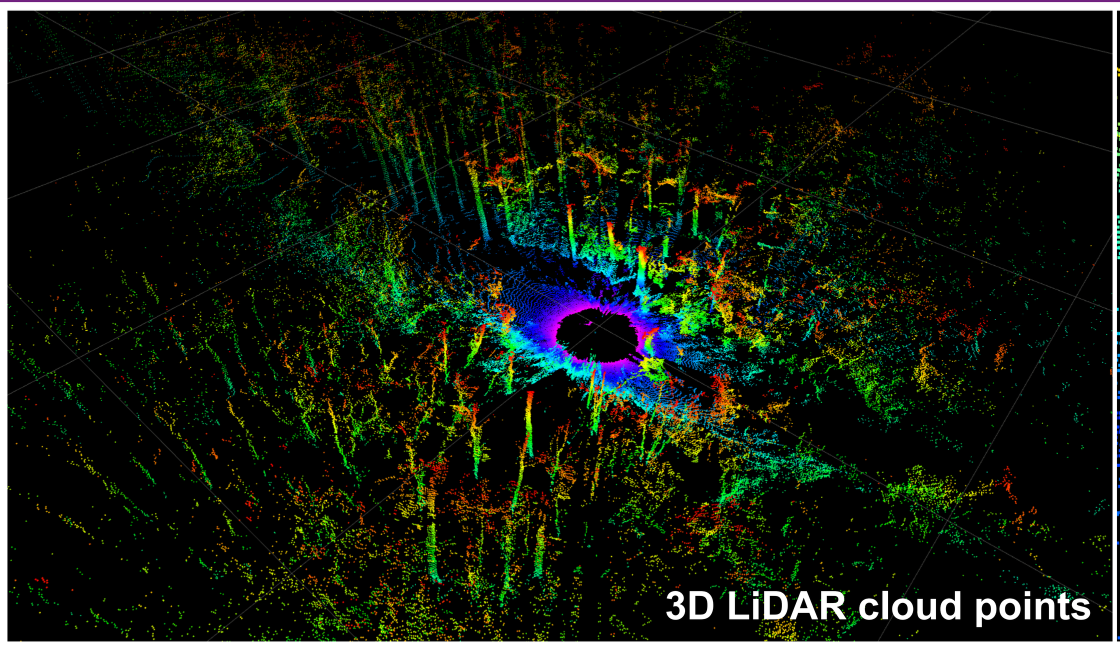
TOMD: A Trail-based Off-road Multimodal Dataset for Traversable Pathway Segmentation under Challenging Illumination Conditions

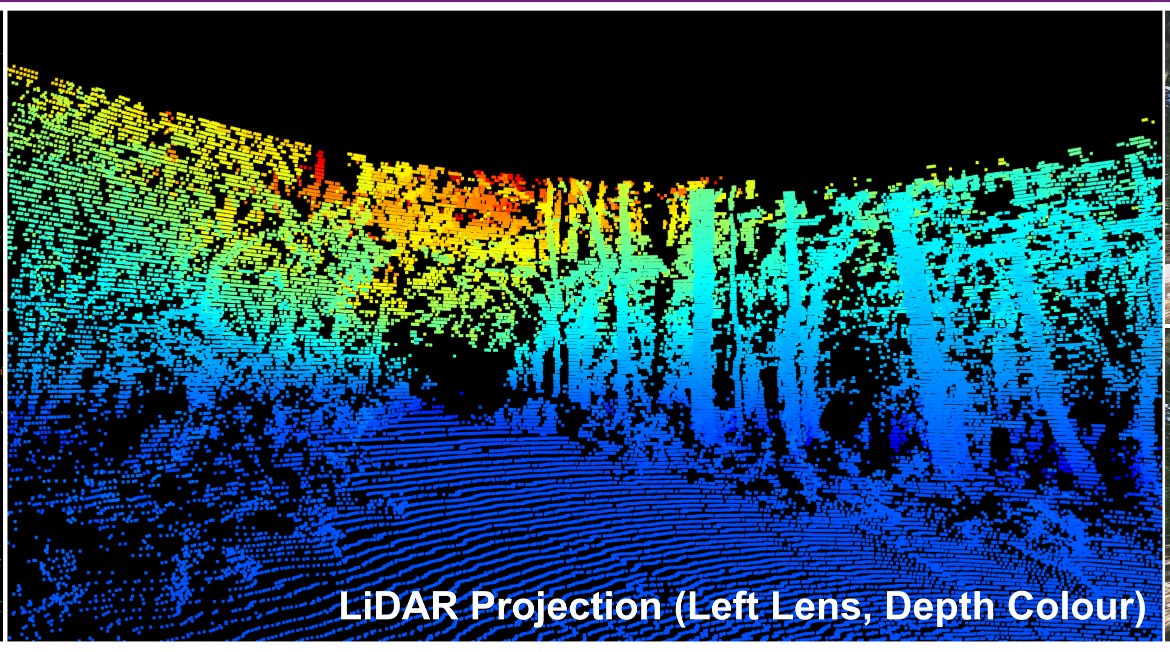
Yixin Sun[1], Li Li[2], Wenke E[1], Amir Atapour-Abarghouei[1], Toby P. Breckon[1]
[1]Durham University, [2]King's College London, UK

Code



Data Collection Platform


3D LiDAR cloud points


LiDAR Projection (Left Lens, Depth Colour)


Left Stereo Image


Right Stereo Image

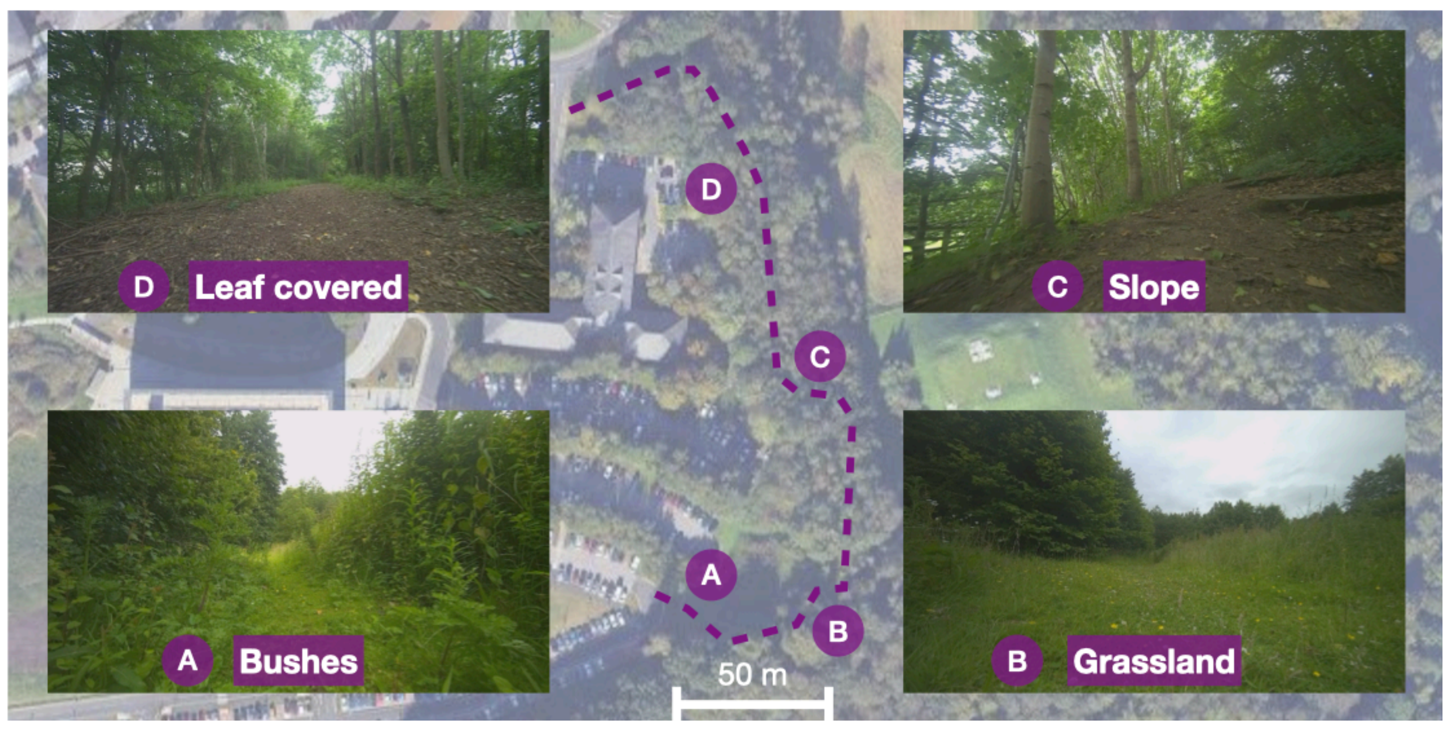
Motivations

- **Real-world relevance**: Autonomous navigation in off-road environments is critical for applications such as search and rescue and disaster response.
- **Insufficient existing datasets**: Most publicly available datasets focus on structured urban roads or wide, vehicle-accessible off-road paths.
- **Trail-specific challenges**: Narrow trails introduce constraints like limited navigable width, dense vegetation, and rapidly changing illumination conditions.

Contributions

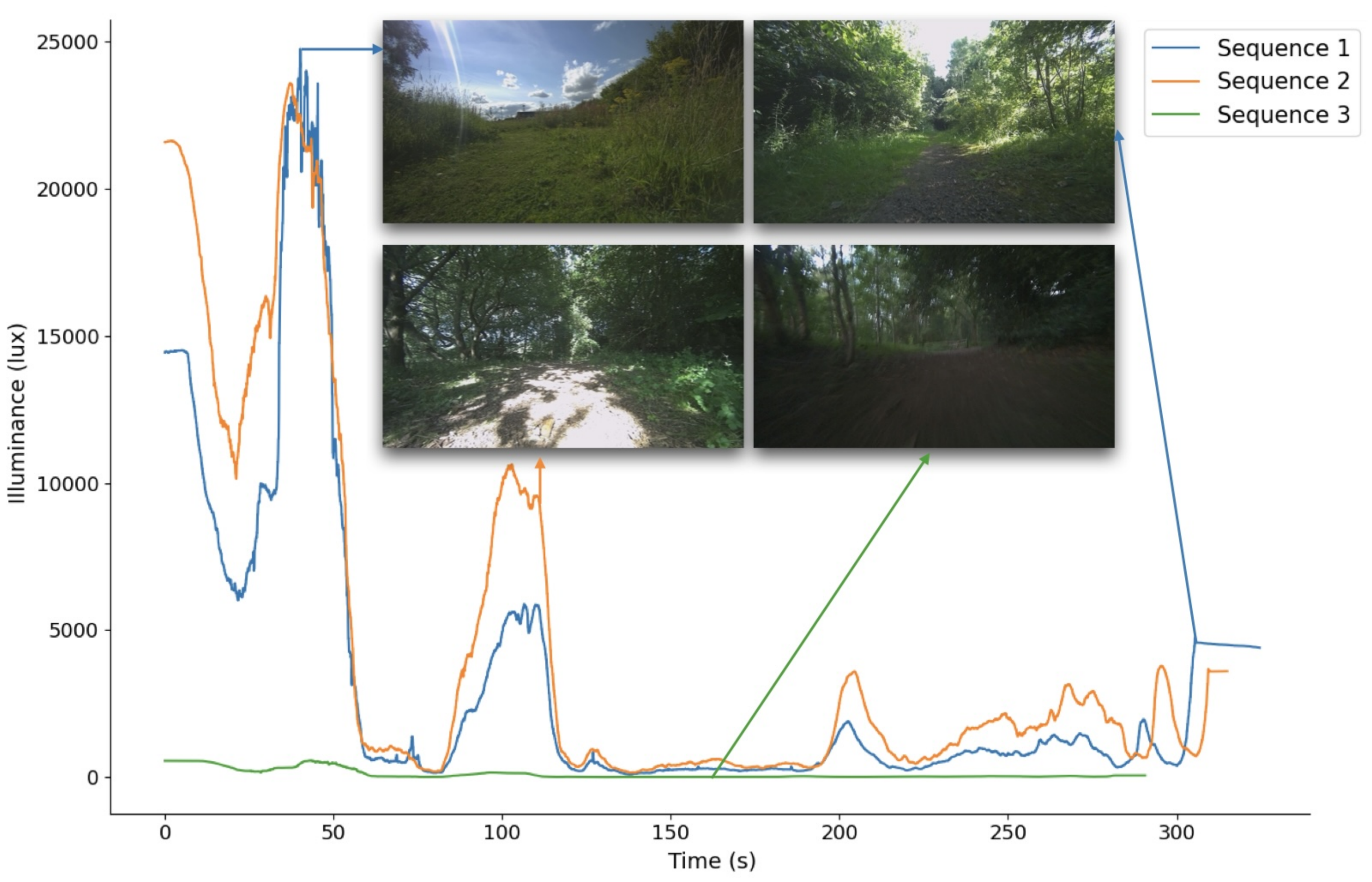
TOMD is the first large-scale multimodal dataset explicitly designed for **unstructured, trail-like** environments, collected using a medium-scale all-terrain robot. Key features include:

- **High-fidelity sensors**: **128**-channel LiDAR, **1920x1080** stereo camera, **real-time kinematic** GNSS, IMU, and Lux meter.
- **Extended modalities**: Ambient illumination measurements, control-level teleoperation commands, and image-level traversability annotations.
- **Accurate Calibration**: Two-stage LiDAR-to-camera calibration strategies for precise sensors alignment.



TOMD collection route route with visual examples of diverse terrain types including bushes, grassland, slope, and leaf-covered trails.

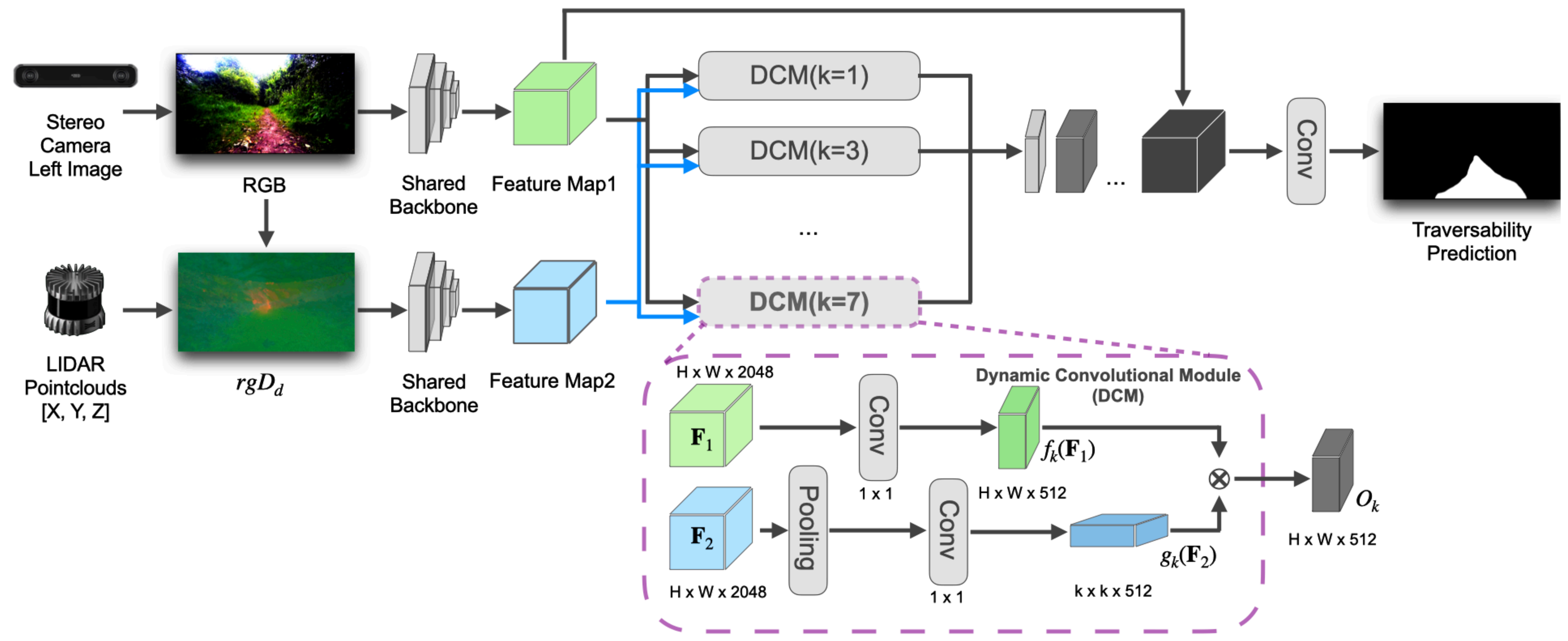
- **High diversity**: Repeated data collection across natural trail routes under diverse terrains and highly variable illumination conditions.



Exemplar TOMD sequences captured under diverse varying ambient illumination levels, including low-light, normal daylight, and high-exposure conditions.

Benchmark Task: Traversable Area Detection

We propose a **dynamic multiscale segmentation framework** with **cross-modality fusion** for accurate traversable pathway detection in trail-like environments. The architecture integrates a **Dynamic Convolutional Module** (DCM) to extract multiscale features in parallel and supports multiple **fusion strategies**—**early** (based on colour chromaticity), **cross**, and **mixed**—to integrate **camera and LiDAR data**. We **systematically evaluate** the impact of each fusion strategy under **varying ambient illumination levels**, highlighting their relative strengths in challenging off-road conditions.



Results

- **Performance Comparison**
  Our **mixed fusion model** surpasses OFF-Net by **+4.46% IoU** and **+2.57% F1-score**, while achieving **25.58 FPS** vs. **15.65 FPS**, showing strong suitability for **real-time off-road applications**.
- **Illumination-Aware Evaluation**
  - **Low light**: RGB-only inputs suffer from underexposure; **depth** provides stable cues. **Dense-depth fusion** notably improves accuracy.
  - **Medium light**: Fusion yields marginal gains, as RGB already captures key features.
  - **High light**: Depth noise slightly affects some fusion strategies, but **early** and **mixed fusion** remain robust.

(a) low (0–100 lux)

Data fusion	Input 1	Input 2	Accuracy (%)	IoU (%)	F1 score (%)
N/A	*RGB*	*RGB*	91.06	80.13	88.46
Early	rgD_s	rgD_s	94.51	86.65	92.36
Early	rgD_d	rgD_d	94.65	87.02	92.72
Cross	*RGB*	D_s	95.49	89.27	94.15
Cross	*RGB*	D_d	**95.78**	**89.99**	**94.57**
Mixed	*RGB*	rgD_s	95.32	89.02	94.01
Mixed	*RGB*	rgD_d	<u>95.62</u>	<u>89.62</u>	<u>94.37</u>

(b) medium (100–10,000 lux)

Data fusion	Input 1	Input 2	Accuracy (%)	IoU (%)	F1 score (%)
N/A	*RGB*	*RGB*	96.71	91.71	95.36
Early	rgD_s	rgD_s	96.54	90.95	94.67
Early	rgD_d	rgD_d	95.77	89.88	94.40
Cross	*RGB*	D_s	**97.61**	<u>93.71</u>	<u>96.56</u>
Cross	*RGB*	D_d	<u>97.60</u>	**93.79**	**96.63**
Mixed	*RGB*	rgD_s	97.24	92.87	96.11
Mixed	*RGB*	rgD_d	97.19	93.05	96.17

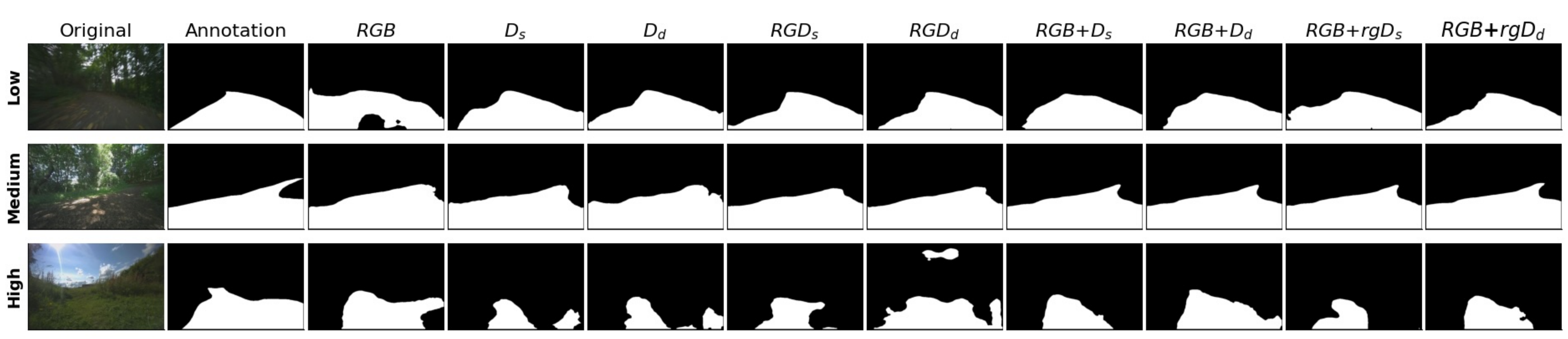
(c) high (>10,000 lux)

Data fusion	Input 1	Input 2	Accuracy (%)	IoU (%)	F1 score (%)
N/A	*RGB*	*RGB*	92.68	79.58	79.58
Early	rgD_s	rgD_s	91.09	75.34	84.33
Early	rgD_d	rgD_d	<u>92.71</u>	<u>79.88</u>	<u>87.95</u>
Cross	*RGB*	D_s	90.88	73.19	80.60
Cross	*RGB*	D_d	91.52	75.40	83.05
Mixed	*RGB*	rgD_s	91.50	75.48	83.11
Mixed	*RGB*	rgD_d	**93.76**	**82.01**	**88.90**

(d) Summary

Data fusion	Input 1	Input 2	Accuracy (%)	IoU (%)	F1 score (%)
N/A	*RGB*	*RGB*	94.18	85.47	89.30
N/A	D_s	D_s	94.74	86.10	91.72
N/A	D_d	D_d	94.91	86.52	92.09
Early	rgD_s	rgD_s	94.53	85.59	91.25
Early	rgD_d	rgD_d	95.07	87.18	<u>92.59</u>
Cross	*RGB*	D_s	95.30	87.18	91.72
Cross	*RGB*	D_d	<u>95.41</u>	<u>87.57</u>	92.25
Mixed	*RGB*	rgD_s	95.24	87.23	92.09
Mixed	*RGB*	rgD_d	**95.85**	**89.16**	**94.27**

Performance comparison under various ambient illumination levels: **bold** / <u>underlined</u> = **best** / <u>2nd best</u> performance.



Traversable area predictions under three ambient illumination levels: **low** (22.80 lx), **medium** (1512.30 lx), and **high** (45,445.10 lx) for differing input modalities: **-D** = LiDAR depth; D_s and D_d indicate sparse (raw LiDAR) and dense (after depth completion), respectively. **RGB** refers to the original colour channels, while **rg** = red and green chromaticity channels, (generated through an early fusion strategy that leverages colour chromaticity).