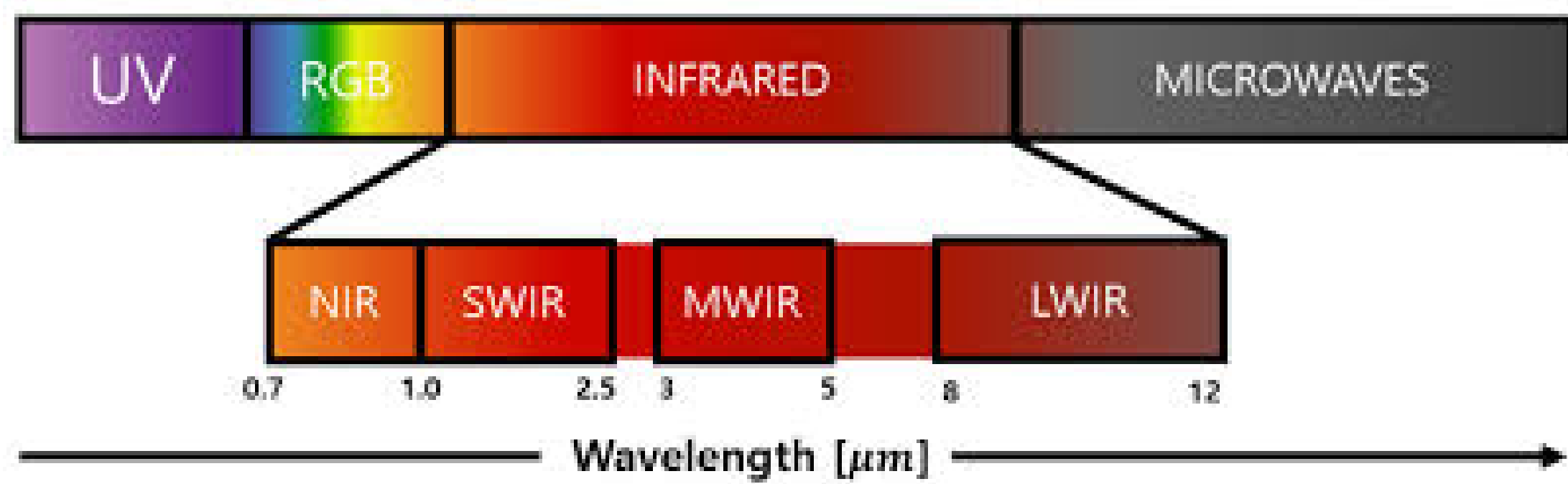


Introduction

RGB cameras dominate autonomous driving yet sometimes fail in adverse weather conditions. SWIR (0.9–1.7 μm) gives us advantages: (1) Penetrates fog, rain, atmospheric scattering. (2) Reveals material reflectance invisible to RGB. (3) Robust to glare and overexposure (4) Maintains contrast on snow-covered surfaces



Source: [5]



Fig 3: SWIR before preprocessing (left) and after normalization (right)

Dataset

RGB And SWIR Multispectral Driving Dataset [1] — first large-scale public synchronized RGB–SWIR driving dataset.

| Property | Value |
|---------------------|----------------------------|
| Total image pair | 100,000 |
| Distance / Duration | 163.3 km / 8.5 h |
| RGB sensor | FLIR GS3 2048×1536 |
| SWIR sensor | CREVIS HG-A130SW 1296×1032 |

Table 1: RASMD Dataset

SWIR Preprocessing: Adaptive threshold at $\mu + 2.58\sigma$ (99% CI) + linear normalization to [0, 255] removes overexposed regions [3].



Fig 3: SWIR before preprocessing (left) and after normalization (right)

Detection & Domain Adaptation

Models Used:

- **YOLOv8x:** CNN single-stage detector [2]
- **RF-DETR:** Transformer-based detector [4]

Fine-tuning strategies: Pre-trained, RGB-only, SWIR-only, Mix of RGB & SWIR (50/50), Synthetic Aug. (20% Fake + 80% Real SWIR), respective weather-specialist RGB, SWIR, and Mix.

| Condition | Modality | YOLOv8 mAP \uparrow | | | RF-DETR mAP \uparrow | | |
|-----------|----------|-----------------------|---------------|---------------|------------------------|---------------|---------------|
| | | RGB FT | SWIR FT | Mix FT | RGB FT | SWIR FT | Mix FT |
| Sunny | RGB | 0.4095 | 0.3116 | 0.3829 | 0.4423 | 0.3662 | 0.4245 |
| | SWIR | 0.2298 | 0.3323 | 0.3380 | 0.3064 | 0.3762 | 0.3470 |
| Cloudy | RGB | 0.4211 | 0.2693 | 0.3944 | 0.4597 | 0.3637 | 0.4173 |
| | SWIR | 0.1792 | 0.2892 | 0.2779 | 0.2808 | 0.3367 | 0.3162 |
| Rainy | RGB | 0.4647 | 0.3026 | 0.4498 | 0.4920 | 0.3867 | 0.4959 |
| | SWIR | 0.2930 | 0.4265 | 0.4113 | 0.3790 | 0.4473 | 0.4356 |
| Snowy | RGB | 0.2125 | 0.1357 | 0.2107 | 0.2705 | 0.2053 | 0.2580 |
| | SWIR | 0.1459 | 0.2717 | 0.2972 | 0.2453 | 0.2705 | 0.2757 |
| Combined | RGB | 0.3962 | 0.2649 | 0.3681 | 0.4374 | 0.3381 | 0.4105 |
| | SWIR | 0.2151 | 0.3279 | 0.3255 | 0.3045 | 0.3651 | 0.3432 |

Table 2: Selected Results

Key Findings:

- Models fine-tuned on the matching modality consistently achieve the highest mAP, Mix fine-tuning gives a balanced strategy and always near the higher score.
- RF-DETR achieves higher scores across all conditions and modalities compared to YOLO, offering better detections.
- In most of the cases aggregated, RGB provides us better scores except snowy where SWIR dominates by a small margin.

SWIR Use-Case Identification: Segregation Methods

To identify when SWIR outperforms RGB across the full 100,000-image dataset (unannotated), three segregation methods developed. Each assigns frames to one of four categories: **RGB-Better**, **SWIR-Better**, **Equal**, or **Discarded** (no decision can be made).

- **M1 (Any-Vote Sorting):** A neutral dataset-wide method. Each of three model groups (RF-DETR, YOLO-MIX, YOLO-Specialist) casts a vote; a single "Better" vote is sufficient to assign superiority. Frames with conflicting votes are discarded.
- **M2 (Majority Consensus):** Requires at least 2-of-3 model agreement for a superiority label, reducing noise.
- **M3 (Strict Unanimous Consensus):** Requires all three models to agree (3/3).

| Method | RGB Better | SWIR Better | Equal | Discarded |
|--------------------|------------|-------------|-------|-----------|
| M1 (Any Vote) | 14.7% | 10.4% | 13.4% | 61.5% |
| M2 (Majority Vote) | 8.6% | 5.0% | 25.0% | 61.5% |
| M3 (Unanimous) | 3.7% | 1.7% | 33.2% | 61.5% |

Table 3: Frame distribution across segregation methods ($N = 99,481$)

SWIR Advantage: Use-Case Analysis

Four conditions were identified from the methods where SWIR consistently and measurably outperforms RGB. The dominant pattern: RGB maintains competitive precision but suffers severe recall collapse under adverse optics, while SWIR preserves structural contrast and detection sensitivity.

1. Hidden Condition (shadow / high contrast)



2. Very-Far Condition (long-range detection)



3. Glare Condition (tunnel exit / sun saturation)



4. Water-Window (rain on windshield)

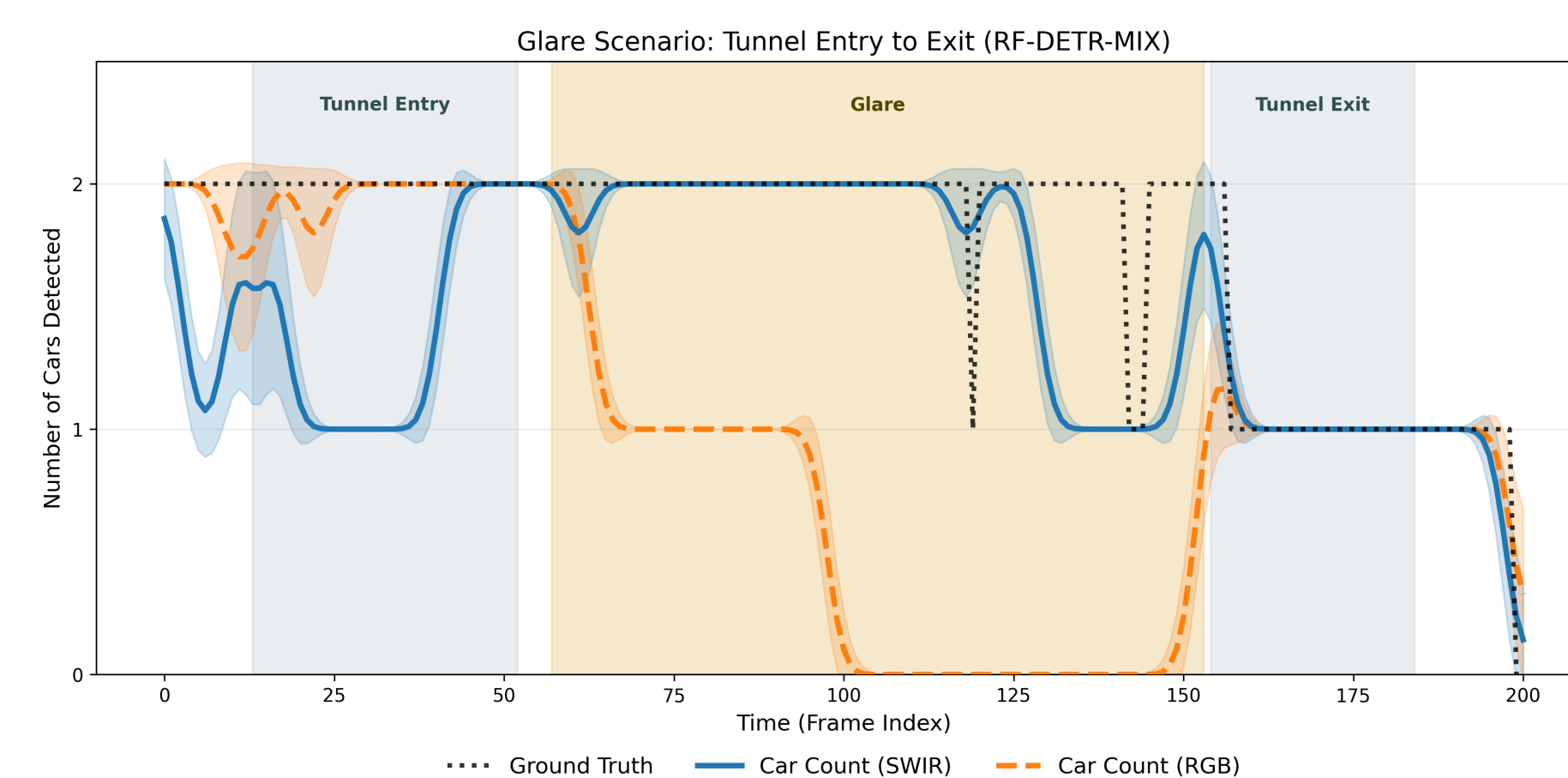
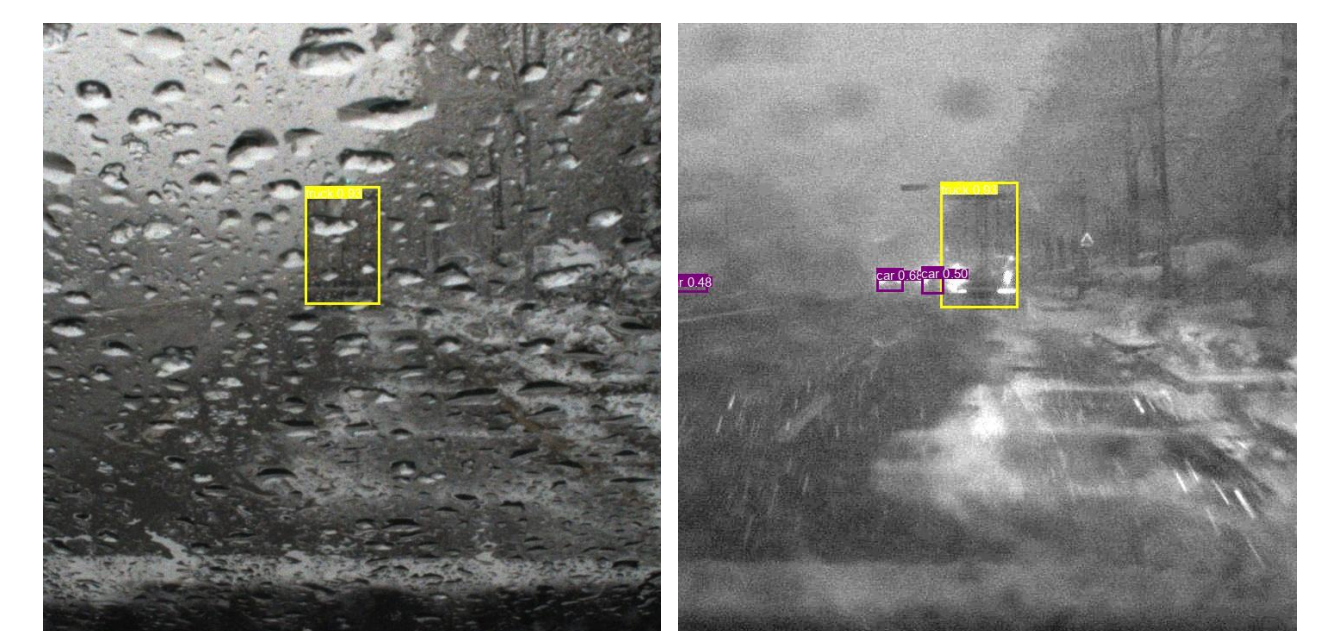


Fig 4: Frame-by-frame analysis of a car going through a tunnel - SWIR dominates over RGB when glare hits.

| Condition | Recall \uparrow | | F-1 \uparrow | | Binary AP \uparrow | |
|--------------|-------------------|--------------|----------------|--------------|----------------------|--------------|
| | RGB | SWIR | RGB | SWIR | RGB | SWIR |
| Hidden | 0.763 | 0.806 | 0.833 | 0.864 | 0.542 | 0.621 |
| Very-Far | 0.729 | 0.789 | 0.811 | 0.852 | 0.503 | 0.572 |
| Glare | 0.332 | 0.720 | 0.475 | 0.808 | 0.195 | 0.384 |
| Water-Window | 0.646 | 0.800 | 0.747 | 0.864 | 0.358 | 0.546 |

Table 4: RGB vs. SWIR on identified use-case conditions after annotations (RF-DETR-Mix).

Conclusions

- **SWIR complements, not replaces, RGB.** RGB leads under nominal conditions; SWIR provides critical redundancy when visible-spectrum sensing degrades.
- **RF-DETR outperforms YOLOv8** across all conditions and modalities, driven by consistently higher recall — essential in safety-critical perception.
- **Synthetic SWIR is viable.** Pix2PixHD Fake+Real fine-tuning closely matches real-SWIR performance, addressing annotation scarcity.
- **SWIR advantage is recall-driven.** Across all four use-cases, precision stays comparable while SWIR recall dramatically outperforms RGB.
- **Future work:** RGB–SWIR sensor fusion and dynamic modality weighting.

References

- [1] Youngwan Jin et al. In: *Information Fusion* (2026).
- [2] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. 2023.
- [3] Rohan Mehra et al. In: *ICCV Workshops*. 2025.
- [4] Pierre Robicheaux et al. In: *arXiv* (2025).
- [5] Hanbin Song et al. In: *Applied Sciences* (2024).