

## RESEARCH ARTICLE

# RecNet: Reinforcement Common Feature Mapping Network for Fast Template Matching in Visible-LWIR Images

DONYUNG KIM<sup>1</sup>, SEUNGEON LEE<sup>1</sup>, INHO PARK<sup>2</sup>, GEONJONG KIM<sup>2</sup>,  
AND SUNGHO KIM<sup>1</sup>

<sup>1</sup>Department of Electronic Engineering, Yeungnam University, Gyeongsan-si 38541, South Korea  
<sup>2</sup>Hanwha Systems, Seoul 04541, South Korea

Corresponding author: Sungcho Kim (sunghokim@yu.ac.kr)

This work was supported by Korea Research Institute for Defense Technology Planning and Advancement (KRIT) Grant funded by Defense Acquisition Program Administration (DAPA) (Augmented reality based Intelligent Aiming System, 2022-2025) under Grant KRIT-CT-22-019.

**ABSTRACT** We present RecNet, a novel end-to-end CNN approach for fast template matching in cross-spectral images, addressing nonlinear intensity disparities and appearance differences in ground-level imagery through a simple yet effective design. Our key innovation lies in the seamless integration of Zero-mean normalized cross correlation (ZNCC), which is well-validated for handling nonlinear intensity variations, with CNN-based shape difference learning. Unlike aerial imagery, ground-level cross-spectral matching presents unique challenges due to limited common features, which RecNet effectively addresses. Through comprehensive experiments, including layer-wise visualization analysis and comparative studies with pooling layer combinations, we validate our architecture's effectiveness. Experiments on KAIST Pedestrian, Log-Gabor Histogram Descriptor(LGHD), and Road Scene datasets demonstrate RecNet's superior performance and real-time capabilities compared to state-of-the-art methods. Additional evaluations using KAIST nighttime imagery and the M3FD dataset verify RecNet's generalization capabilities and stability across diverse scenes and conditions, while also identifying limitations and future research directions.

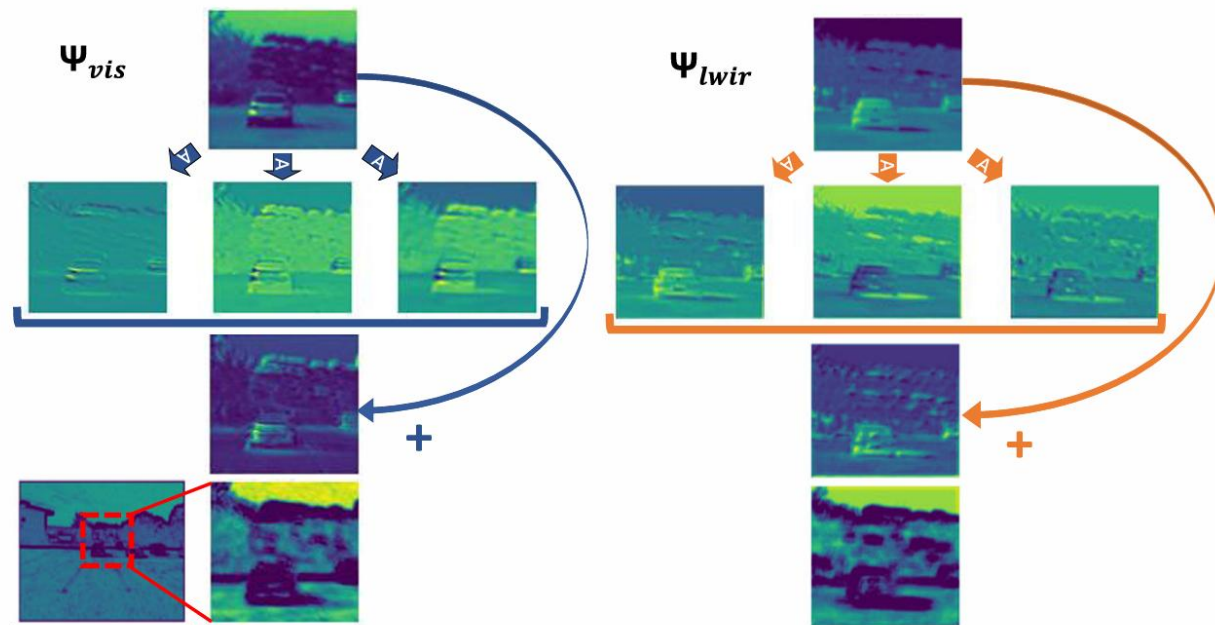
**INDEX TERMS** Visible-LWIR template matching, multi-modality, cross-spectral image processing.

## I. INTRODUCTION

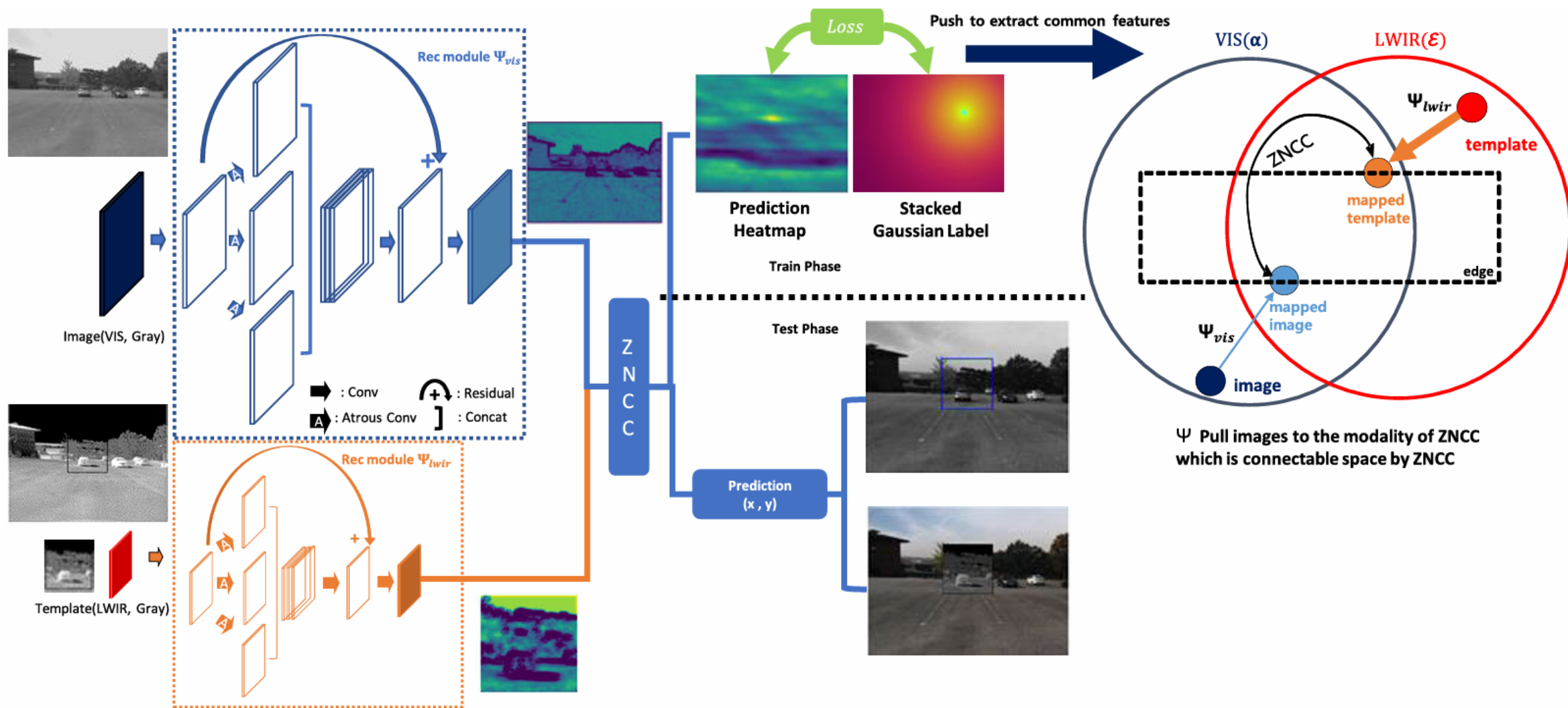
Recent advancements in computer vision have revolutionized various image processing techniques are being used in daily life and industrial settings like [1], [2], and [3]. Computer vision technology is rapidly evolving, with cross-spectral template matching emerging as a crucial research direction for bridging the capabilities of visible and thermal imaging systems. Template matching, a fundamental computer vision operation, involves systematically scanning a larger search image to identify a specified template pattern through comparison processes [4]. With the advent of deep learning, template matching capabilities have considerably improved, enabling more sophisticated and robust pattern recognition

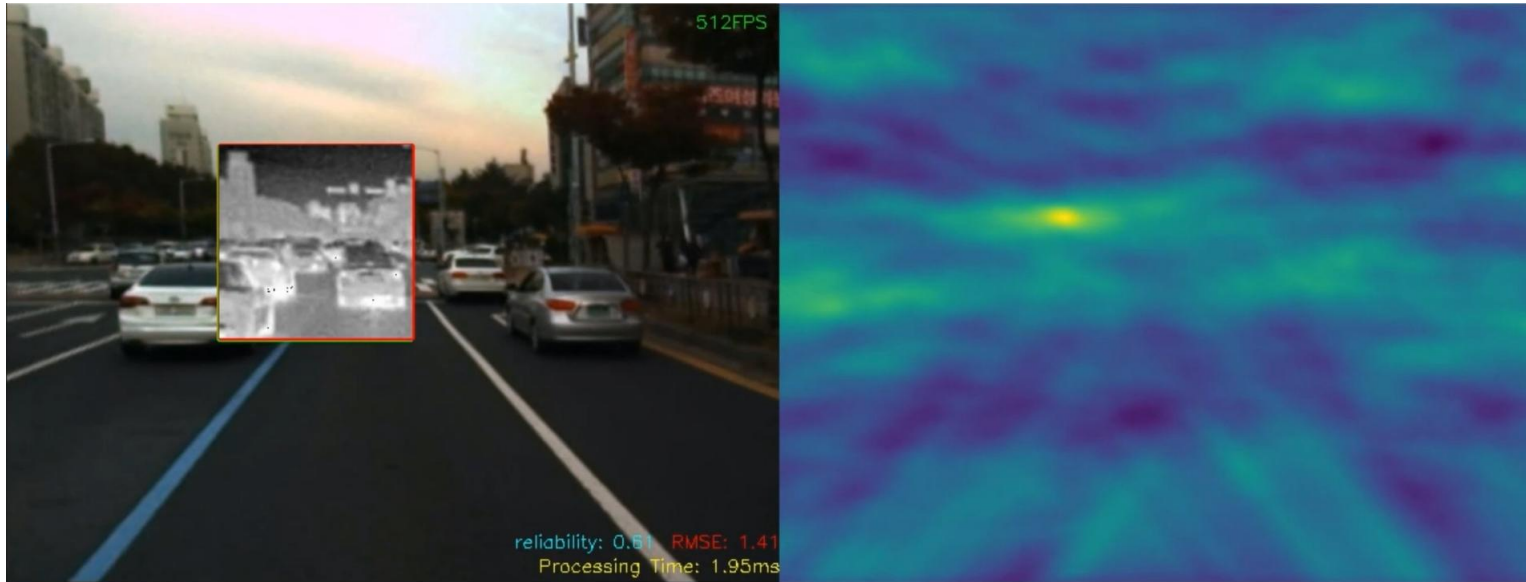
across diverse applications [5]. Long-wave infrared (LWIR) imaging has become increasingly important due to its ability to capture thermal radiation patterns, providing complementary information to traditional visible light imaging [6]. This dual-modality approach is particularly valuable in scenarios requiring robust object detection and recognition under challenging lighting conditions [7]. The integration of visible and LWIR imaging has enabled enhanced perception capabilities, leading to successful applications in autonomous navigation, surveillance systems, and medical imaging [8]. However, cross-spectral template matching between visible and LWIR images presents unique challenges. The fundamental difference in their sensing mechanisms visible images capturing reflected light versus LWIR images detecting thermal emissions—results in significant appearance disparities when representing the same

# 지능형 화기조준경 프로젝트 EO-LWIR 영상 매칭 알고리즘



The associate editor coordinating the review of this manuscript and approving it for publication was Geng-Ming Jiang.





EO-LWIR 영상 매칭 실시간  
매칭 결과 및 히트맵





# IR-Radar 추적 미사일 알고리즘

## 거리 시간별 합성 및 채프 RF·IR 신호 모사 및 추적

김도영<sup>1</sup>, 장오태<sup>2</sup>, 조상호<sup>2</sup>, 김성호<sup>2\*</sup>

영남대학교 로봇공학과<sup>1</sup>, 영남대학교 전자공학과<sup>2</sup>

\* 교신저자

dnfdnfdn06@gmail.com

### 요약

본 논문은 해상환경에서 미사일과 함정의 전장 상황을 레이더, 적외선 신호 모사를 통한 시뮬레이션 속에서 미사일의 합성 추적 모듈 제작에 대한 연구이다. 적외선 신호는 영역제한 기법으로 추적 모듈을 제작하였고 레이더 신호는 선형 칼만 필터를 이용한 추적 모듈을 제작하였다. 레이더추적 모듈과 적외선 추적 모듈을 통합하는 통합 추적 모듈을 제작하여 시뮬레이션에서 사용할 수 있도록 하였다. 추적 모듈과 미사일 제어를 연동시켜 제적이 변화하도록 하였다. 시뮬레이션상에서 배가 채프와 플레이를 전개하는 상황을 시간에 따라 모사하여 함정의 이동경로에 따른 기만 성능을 확인이 가능하도록 하였다. 이를 통해 레이더, 적외선 통합 추적 모듈을 가진 미사일과 함정의 전장 상황을 모사하여 최적 침로 분석 등 다양한 연구의 시험환경을 마련하였다.

### 1. 서론

응용과학이 집약된 현대의 전쟁은 예측하기 어렵고 복잡도가 높아져, 전장 모의환경의 구축과 시뮬레이션 기반 획득을 통한 분석이 필수적이다.[1][2][3] 특히 본 연구에서 모사한 배와 미사일의 전장 상황을 실험하기 위해서는 큰 재원이 소요된다. 이를 대체하기 위해 모의 전장을 통한 개발 및 훈련을 진행하면 경제적 이점이 있는 자명하다. 본 연구에서는 해상환경과 배를 렌더링 한 시뮬레이터 상에서 추적 미사일의 동작을 모사하였다. 유의미한 실험환경을 위해 이동하는 미사일 좌표계를 기준으로 관측되는 각각의 적외선 및 레이더 신호를 모사하였고 통합 추적 모듈을 통해 추적 및 미사일 제에 반영되도록 제작했다.

### 2. 통합 추적 모듈

본 장에서는 적외선 및 레이더 신호 모사에 대하여 설명한 후 미사일 추적을 모사하기 위한 통합 추적 모듈에 대해 다룬다.

### 2.1 적외선 및 레이더 신호 모사

적외선 신호를 모사하기 위해 배, 바다, 플레이를 IR 신호 값으로 렌더링 된 데이터를 이용하였다. 공간 좌표상 표현된 값들을 MATLAB 내장함수 view를 이용하여 미사일 시점의 화면을 연출했다. 레이

더 신호 모사는 실제 관측 데이터를 기반으로 제작된 레이더 신호 계산 함수를 사용하였다. 제작 함수는 배, 채프, 미사일의 좌표를 입력으로 받아 신호 값과 위치를 출력한다. 레이더 신호 모사에서 채프는 플레이와 동일한 좌표를 사용하였다.

### 2.2 적외선 추적 모듈

레이더 추적 모듈은 미사일 시점의 IR 영상을 최대 최소 정규화, 필터링, 영역제한 함수를 통해 최대 영역을 추적하여 관심 영역과 예측 영역을 출력하는 모듈이다.

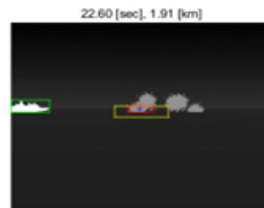
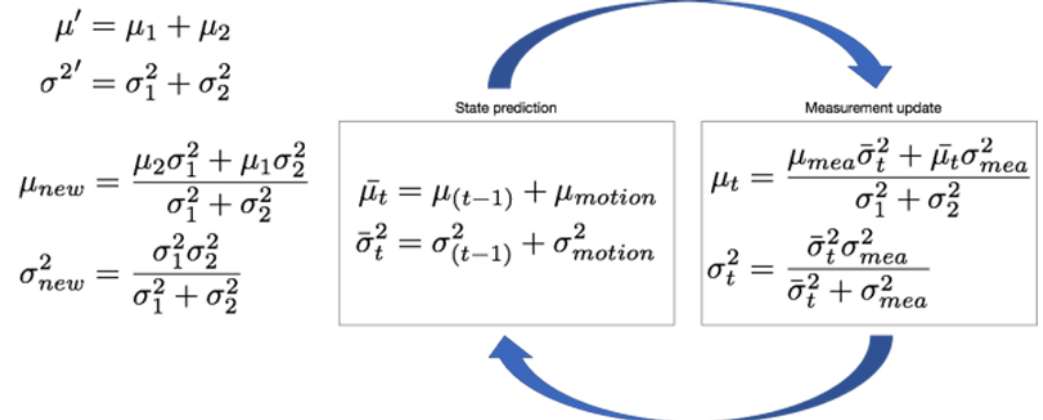


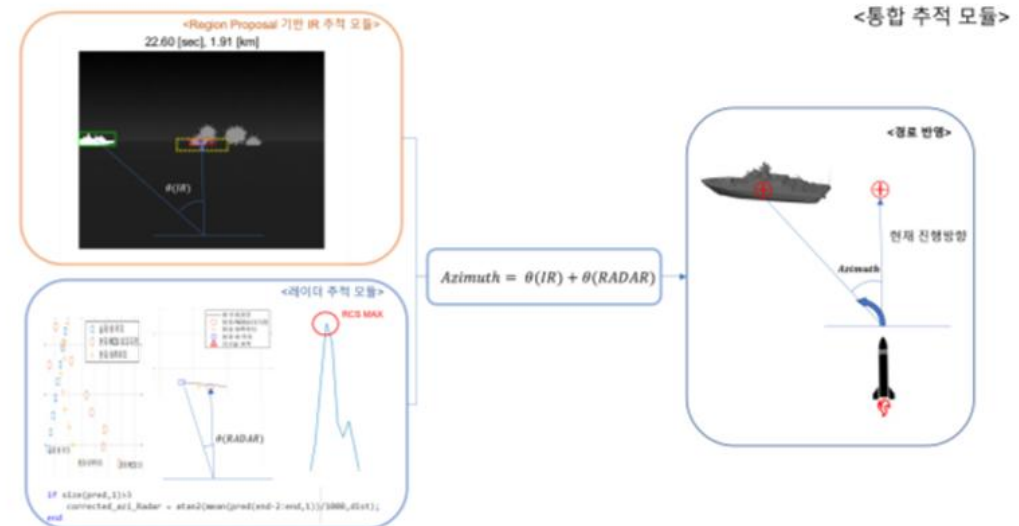
그림 1. IR 추적 모듈

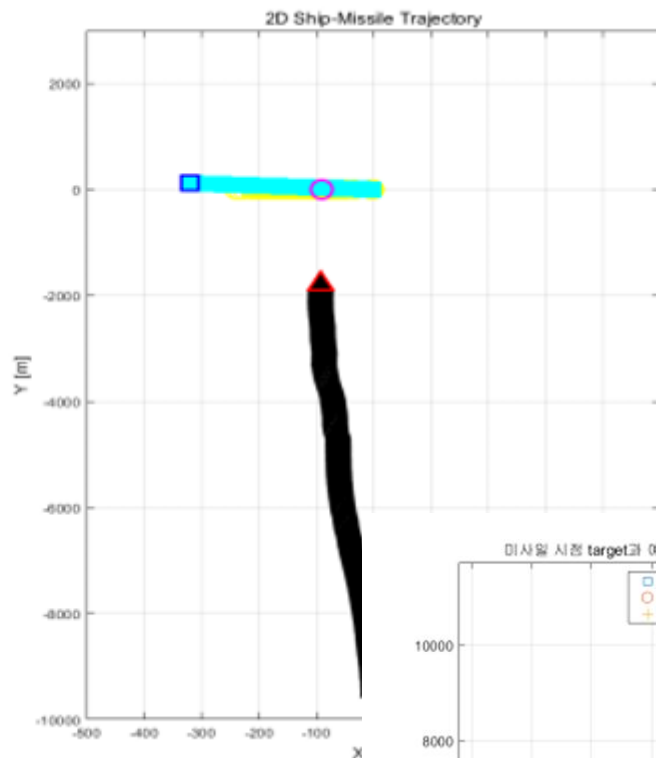
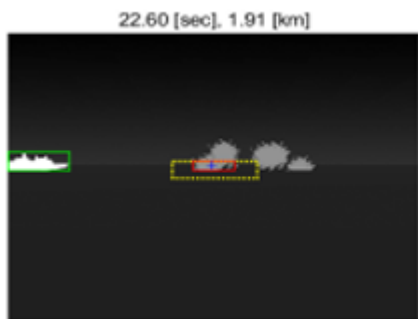
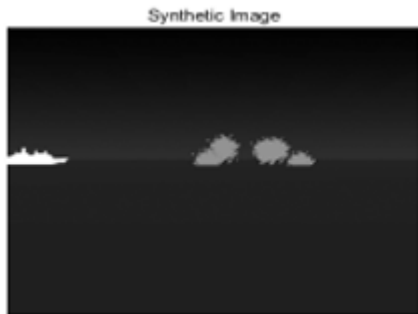
### 2.3 레이더 추적 모듈

레이더 추적 모듈은 그림 4의 세번째 플롯과 같이 레이더 신호가 최대 값 위치에서 두드러지는 특성에 따라 최대 값의 좌표를 입력으로 사용하였다.

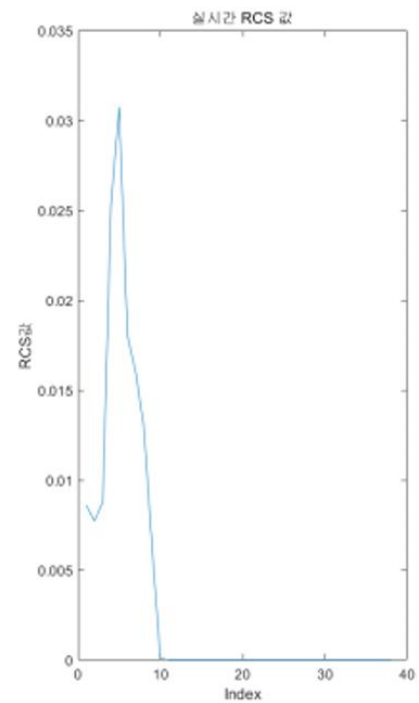
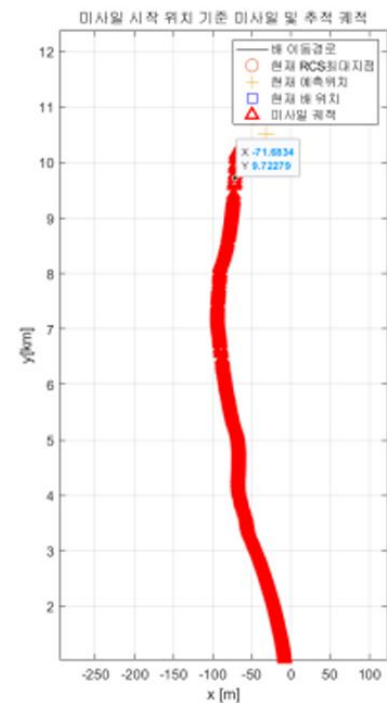
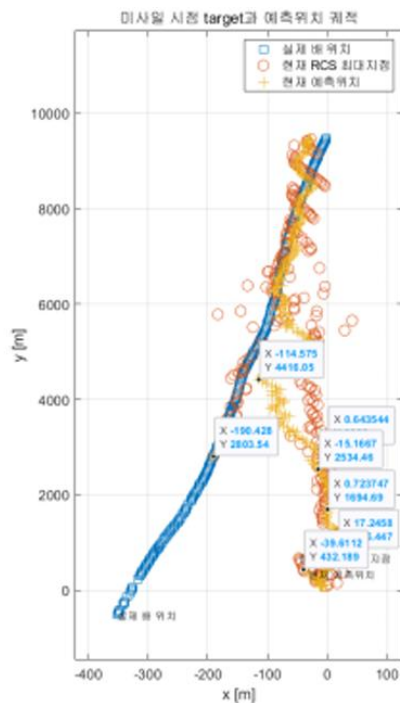


## 수식 1. 선형 칼만필터





플레어 기만상황 결과



IR 신뢰도 통합 추적 모듈 결과

# 영상기반 활동 모니터링 시스템

단안 카메라 위치정보 추출 시스템

CCTV #1



Digital Twin



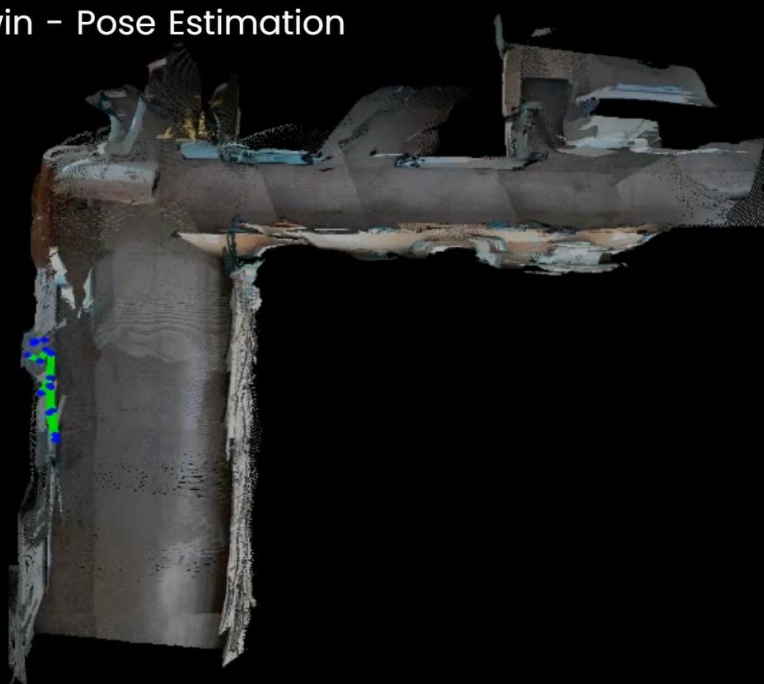
YOLO-V8, DuST3R(3D Map), 3D-Net, 바닥 특징 점 기반 맵핑



CCTV #1



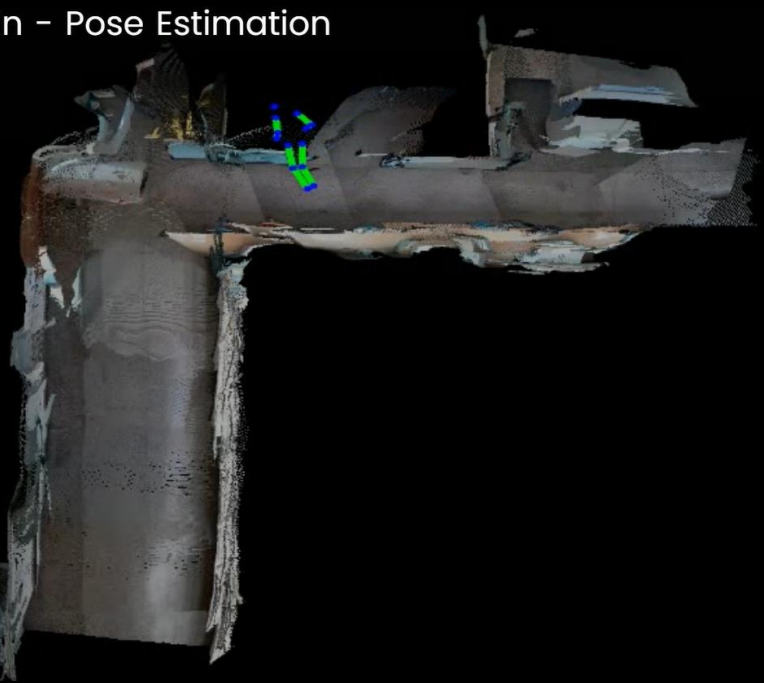
Digital Twin - Pose Estimation



CCTV #2



Digital Twin - Pose Estimation



## 멀티 View 통합 위치 및 자세 추정 시스템

- Yolo-V8, DuST3R(3D Map)

- 포인트 맵을 직접적인 위치 추정(Proposed)

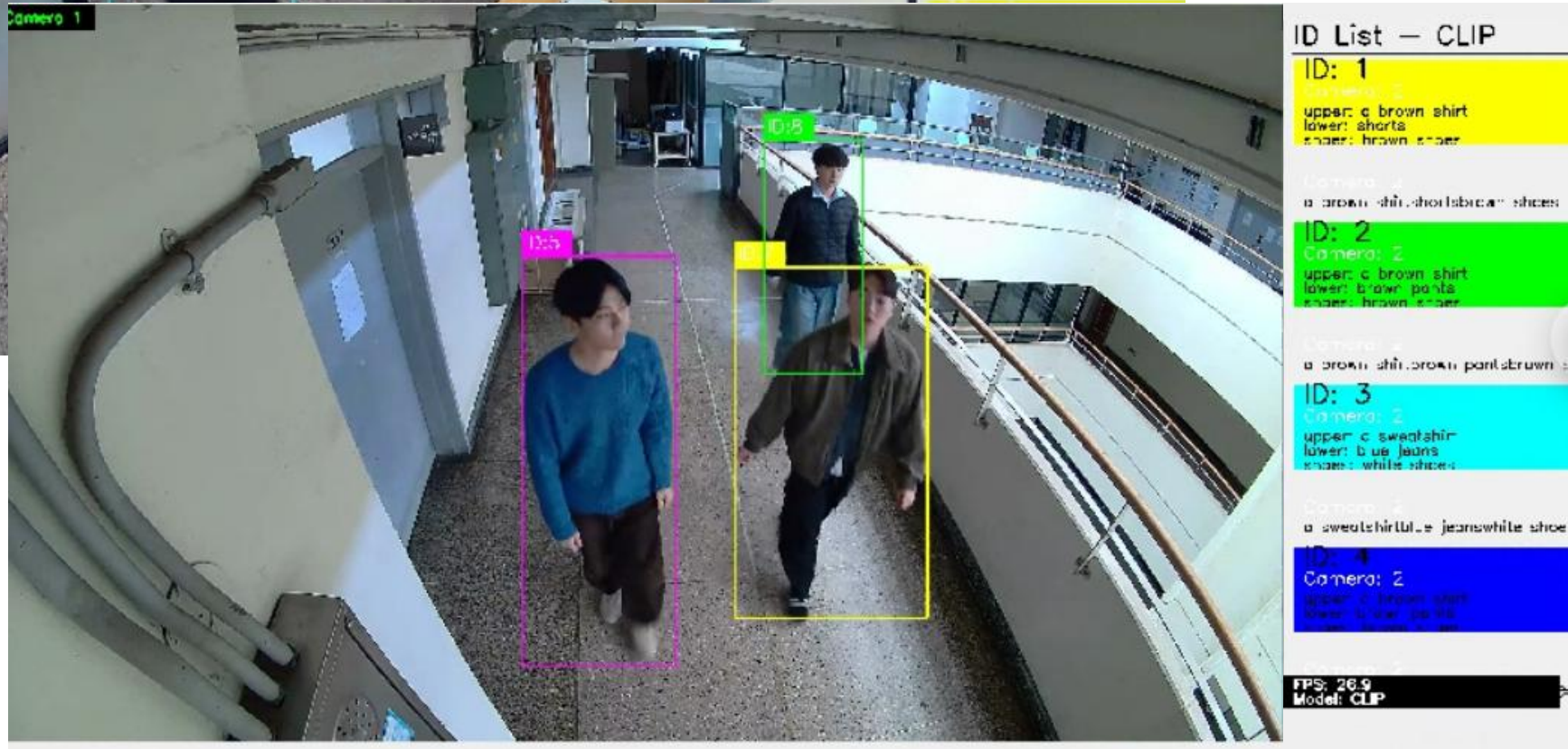




Re-ID(DINO-V2)

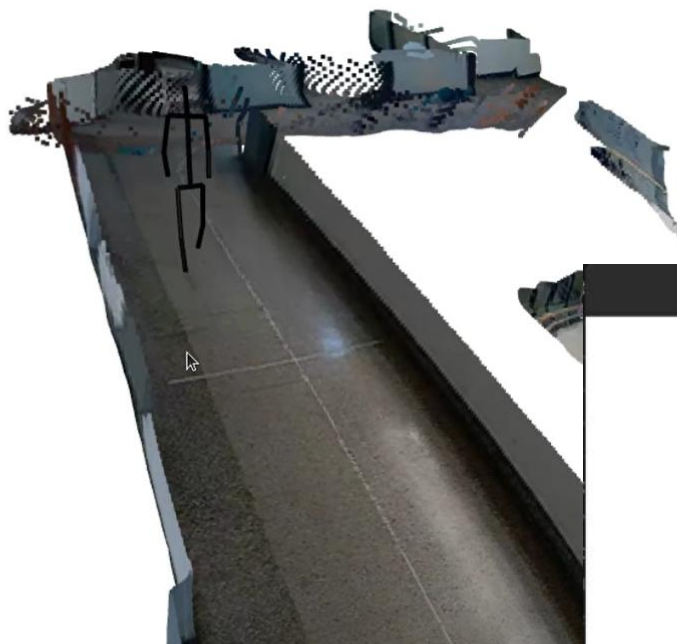
## Image-Text

Re-ID(CLIP: 상의, 하의, 신발에 대한 설명자로 사용)





/home/roadgood/Desktop/2025/digital-twin/marauder\_clip/Marauders\_Map.py

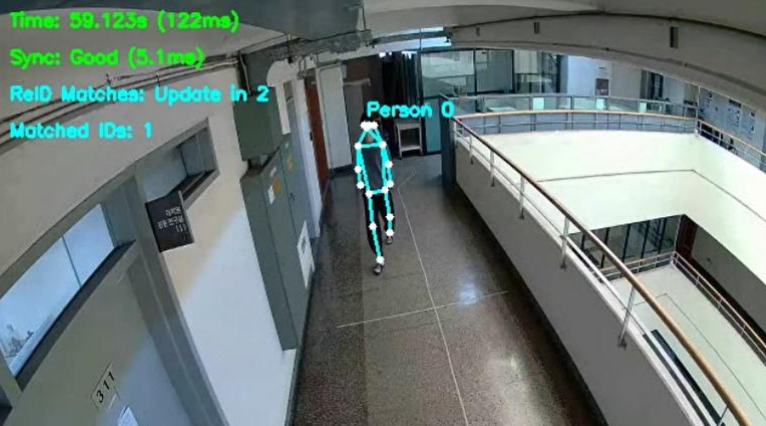


Time: 59.123s (122ms)

Sync: Good (5.1ms)

ReID Matches: Update in 2

Matched IDs: 1



Yolo-V8, DuST3R(3D Map)  
포인트 맵을 이용한 맵핑,  
3D Pose 추정 방식 -

/home/roadgood/Desktop/2025/digital-twin/Marauder\_Multi/Marauders\_Map.py

Multi

2025-03-24 15:56:59.555 INFO ReID matched 1 person



View 0 - 2D Pose Detection



Time: 101.019s (157ms)

Sync: Fair (24.6ms)

ReID Matches: 0

Matched IDs: 0



(x=80, y=52) ~ R:255 G:255 B:0

View 1 - 2D Pose Detection



Time: 100.995s (984ms)

Sync: Fair (24.6ms)

ReID Matches: 0

Matched IDs: 0



(x=123, y=28) ~ R:0 G:255 B:0

# Multi modal 3D Reconstruction Project

## Generation of 3D LWIR thermal maps based on deep learning SLAM: feasibility and evaluation

Donyung Kim<sup>a</sup>, Sungho Kim<sup>\*\*</sup>

<sup>a</sup>Department of Electronic Engineering, Yeungnam University, South Korea;

\* Corresponding author. E-mail addresses: dnfdnfd106@gmail.com, sunghokim@yu.ac.kr (Sungho Kim)

### ABSTRACT

Machine Learning has played a major role in various applications including Visual Slam and thermal image process. In this paper, we discussed the possibility of generating a thermal map using LWIR images and a deep learning-based visual slam network and the value that the thermal map can create. We summarized the advantages and applicability of various deep learning-based visual slams and confirmed the results of nice slam, which generates the most curious Dense map. In order to apply Visual SLAM technology, time series, scene repetition, and images from various angles for one scene are required. However, most LWIR data sets consist of one shot for each scene or are unidirectional driving data. To solve this, we created a scenario using the LWIR driving dataset and created a repetitive route through repetition. RGB-Depth SLAM Mapping was performed on the constructed data set, and the results were evaluated and the limitations of the current approach were discussed. Finally, we summarized future directions for creating stable 3D thermal maps in indoor and outdoor environments by resolving the limitations.

**Keywords:** Multi-Modal image processing, Vis-Lwir, SLAM, Thermal Map, Deep-Learning, 3D mapping, CNN, Multi-spectral

### 1. INTRODUCTION

Machine learning, particularly deep learning, has played a pivotal role in various applications, including visual simultaneous localization and mapping (SLAM) and thermal image processing [1][2]. This paper explores the feasibility and evaluation of generating 3D thermal maps using long-wave infrared (LWIR) imagery and deep learning-based visual SLAM networks. The importance of 3D thermal mapping cannot be overstated. LWIR sensors often face a trade-off between resolution and field of view (FOV) limited to objective factors such as fewer infrared band materials, difficulty in aberration correction, and the limited number of optical elements, the design of ultra-wide FOV infrared staring imaging system [3]. By generating 3D thermal maps, it becomes possible to predict heat conduction and effectively expand the FOV, thereby creating value. Additionally, recent advancements in deep learning have led to the development of visual SLAM models [4]. For example, Droid-SLAM [5], Nice-SLAM [6], and Nicer-SLAM [7], which have opened up possibilities for their application to LWIR imagery.

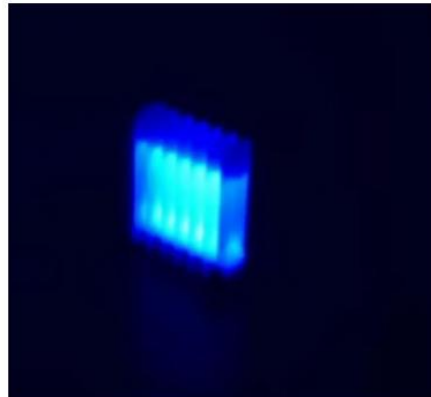
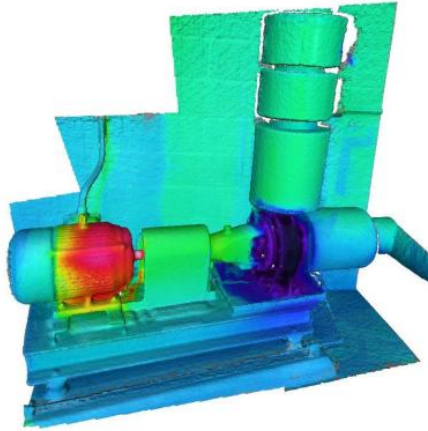
This paper summarizes the characteristics of various deep learning-based visual SLAM models and explores the application of Nice-SLAM, a model capable of generating dense maps, to LWIR data. It outlines the methodology for constructing a scenario-based dataset suitable for SLAM, the results obtained, and an analysis of the limitations. Finally, it proposes future solution directions to achieve stable 3D thermal mapping using deep learning-based visual SLAM techniques.

### 2. RELATED WORKS

3D thermal mapping has been a subject of research for some time, with point clouds method relying on classical registration techniques. This section reviews previous work on 3D thermal mapping models and visual SLAM algorithms like Droid-SLAM, Nice-SLAM

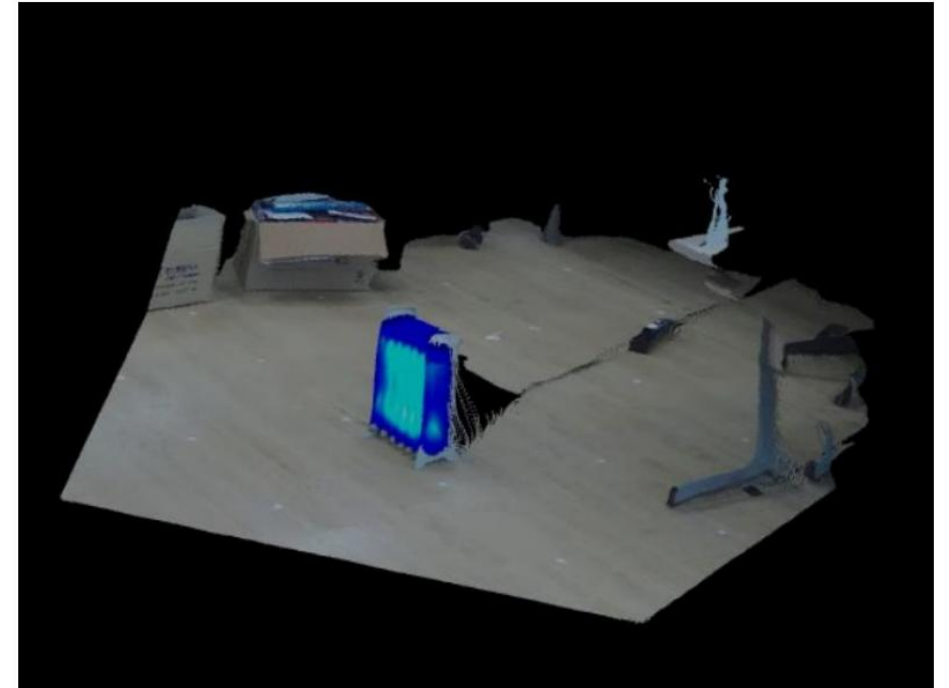
#### 3D Thermal Mapping using point clouds

Borrmann et al. propose a framework for mapping thermal attributes from point clouds onto building facades to generate thermal textures for 3D analysis [8]. They use mobile laser scanning point clouds with thermal infrared intensities and project them onto facades using a nearest neighbor mapping algorithm. The approach "Minimize



## 3D Thermal-Map

- Matching: RecNet(Proposed)
- 3D Map: Dust3R

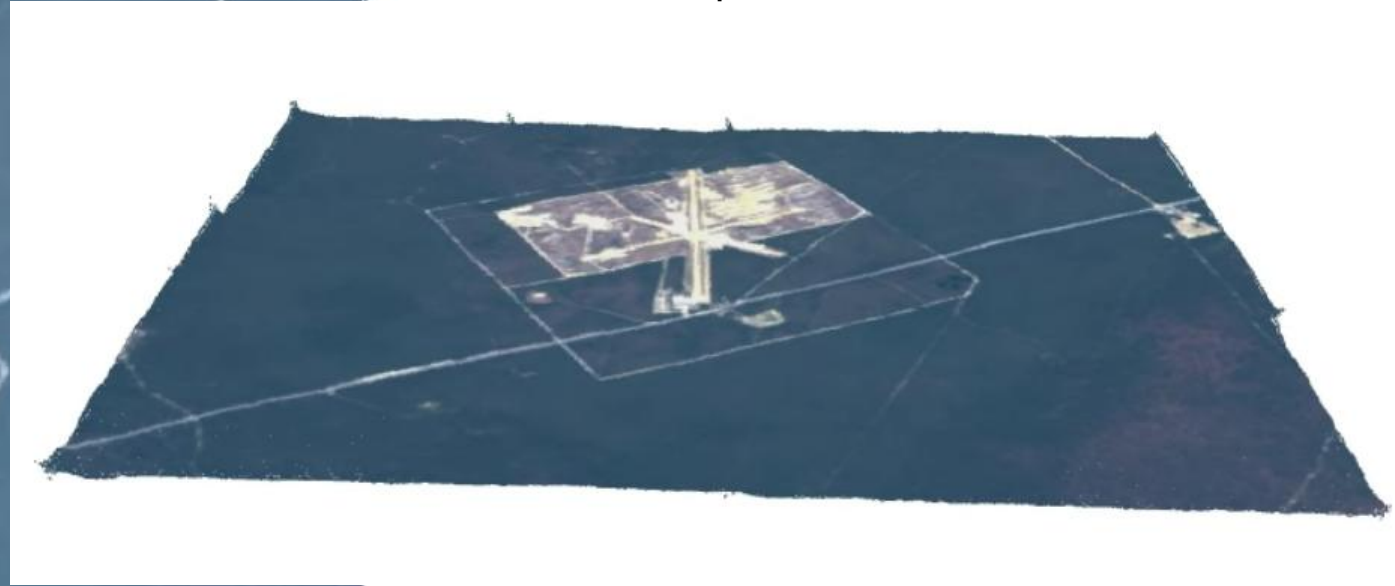


Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications VI, edited by  
Peter J. Schwartz, Benjamin Jensen, Myron E. Hohl, Proc. of SPIE Vol. 13051,  
130510T · © 2024 SPIE · 0277-786X · doi: 10.1117/12.3013365

Proc. of SPIE Vol. 13051 130510T-1

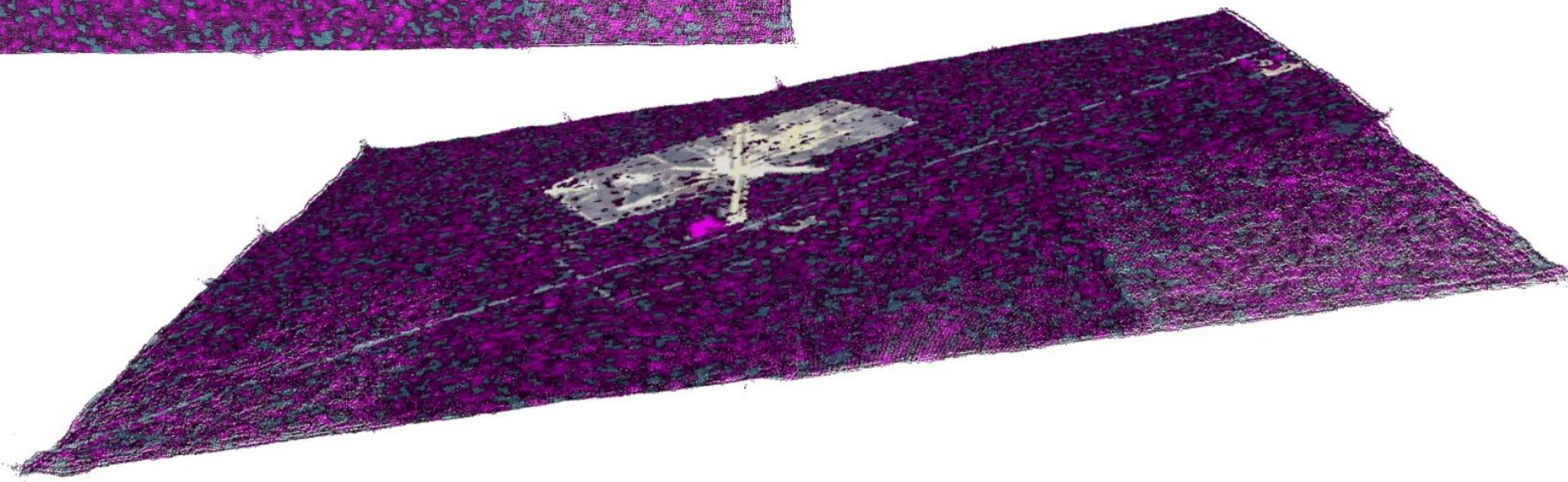
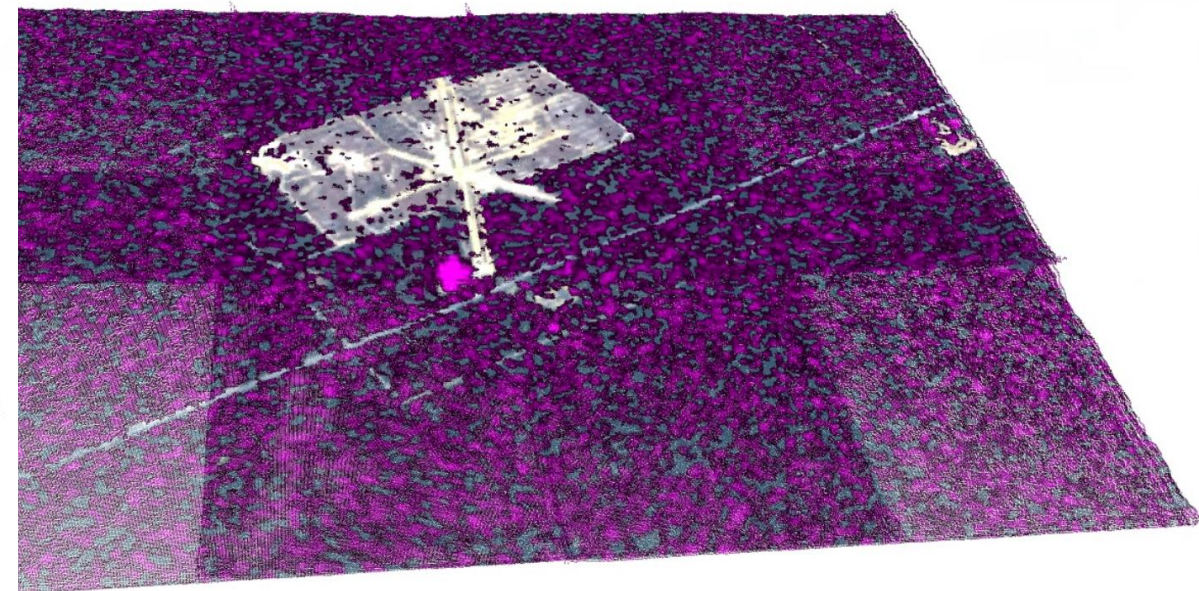
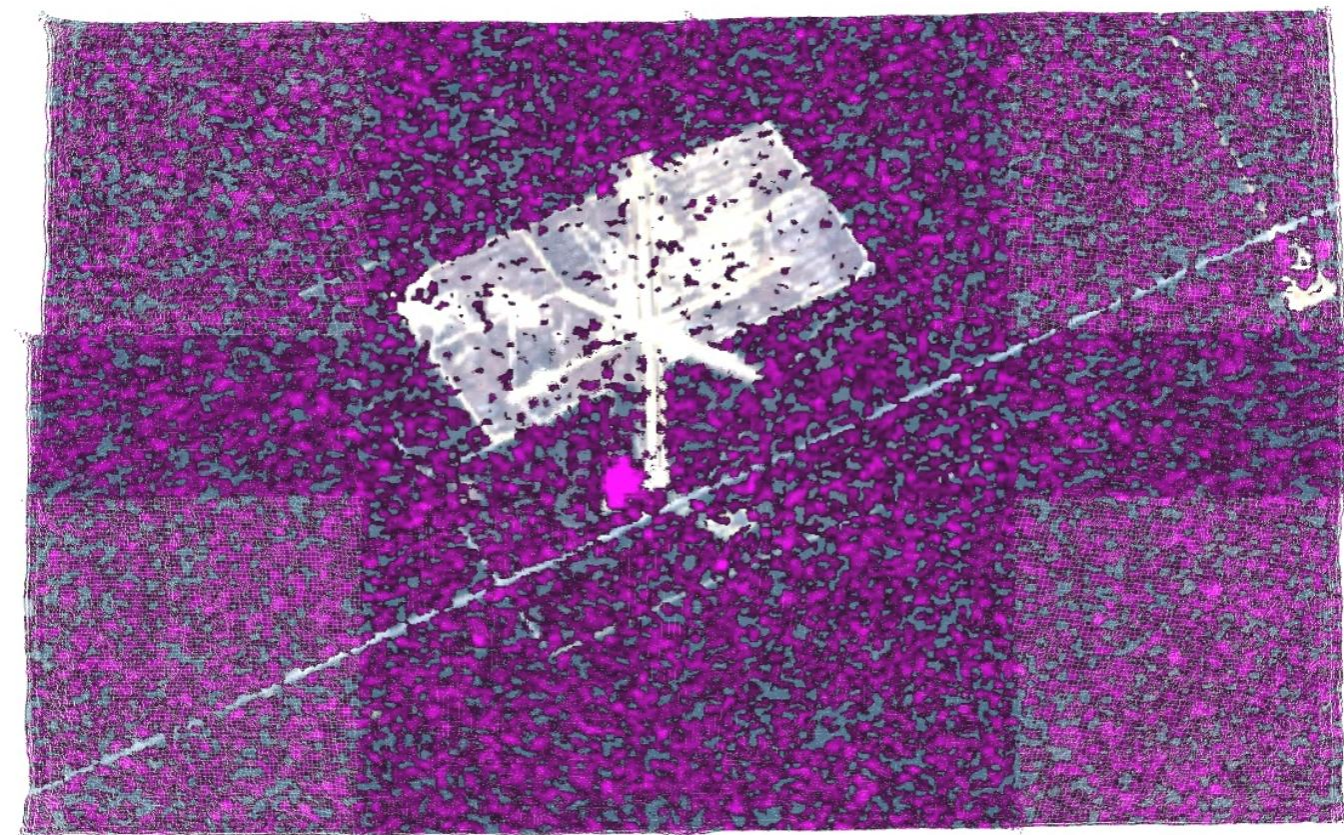


3D SAR Map  
- m1436631 Dataset  
- 3D Map: Dust3R





SAR Threshold 20





SAR Threshold 150

