Check for updates

Journal of KIIT. Vol. 24, No. 10, pp. 137-145, Oct. 31, 2026. pISSN 1598-8619, eISSN 2093-7571
http://dx.doi.org/10.14801/jkiit.2026.24.10.137

초소형 정찰 로봇용 열영상 사람 검출 모델 성능 비교

강전성*[1], 박범준*[2], 윤준원*[3], 송하윤*[4], 정현준**, 서덕현***[1], 윤태준***[2], 김주영***[3]

Performance Comparison of Thermal Human Detection Models for Small Reconnaissance Robots

Jeon-Seong Kang*[1], Beom-Joon Park*[2], Junwon Yoon*[3], Ha-Yoon Song*[4], Hyun-Joon Chung**, Duck-Hyun Suh***[1], Taejoon Yoon***[2], Jooyoung Kim***[3]

This work was supported by Korea Research Institute for defense Technology planning and advancement (KRIT) grant funded by the Korea government (DAPA (Defense Acquisition Program Administration)). (No. KRIT-CT-22-006-005, Control technology for collective operation of military ultra-small ground robots(Contribution rate: 100%))

요 약

초소형 군집 정찰 로봇은 야간·저조도 환경에서 신뢰할 수 있는 열영상 기반 사람 검출 성능이 필수적이다. 본 연구에서는 FLIR Lepton 3.5와 PureThermal2를 사용해 실제 운용 시나리오를 반영한 열영상 데이터셋 5,000장을 구축하고, YOLOv8 계열 모델(n, s, m, l, x)의 사람 검출 성능을 Jetson Orin NX 환경에서 비교하였다. 실험 결과, YOLOv8s와 YOLOv8m은 YOLOv8n보다 높은 검출률과 실시간 추론 속도, 적절한 자원 사용량을 동시에 달성해 소형 로봇용 열영상 사람 검출에 가장 적합한 모델로 확인되었다. 반면 YOLOv8l과 YOLOv8x는 검출률 향상 폭에 비해 연산 자원 부담과 지연이 커 실시간 정찰 임무 적용에는 한계가 있었다. 이러한 결과는 열영상 기반 소형 정찰 로봇에서 경량, 중형 YOLOv8 모델이 성능과 실시간성, 임베디드 자원 제약을 균형 있게 만족하는 최적 해법임을 시사한다.

Abstract

Reliable thermal human detection is essential for small swarm reconnaissance robots operating in low-light and degraded visual environments. We construct a 5,000-image thermal dataset using a FLIR Lepton 3.5 sensor and a PureThermal2 module, and compare the human detection performance of YOLOv8 models (n, s, m, l, x) on a Jetson Orin NX embedded platform. The results show that YOLOv8s and YOLOv8m provide the best trade-off for thermal human detection on small robots, achieving higher detection accuracy than YOLOv8n while maintaining real-time inference speed and acceptable resource usage. In contrast, YOLOv8l and YOLOv8x offer limited accuracy gains at the cost of significantly increased computational load and latency, making them less suitable for real-time reconnaissance missions on resource-constrained platforms. Lightweight and medium YOLOv8 models best balance accuracy, real-time performance, and resource usage for thermal small reconnaissance robots.

Keywords

thermal human detection, embedded vision system, YOLO model, jetson orin nx

* 한국로봇융합연구원 AI휴머노이드본부
- ORCID[1]: https://orcid.org/0000-0002-7061-5794
- ORCID[2]: https://orcid.org/0000-0001-6905-7740
- ORCID[3]: https://orcid.org/0009-0000-1737-5181
- ORCID[4]: https://orcid.org/0009-0007-5935-9797
** 한국로봇융합연구원 AI휴머노이드본부(교신저자)
- ORCID: https://orcid.org/0000-0001-5533-9587
*** LIG Defense & Aerospace
- ORCID[1]: https://orcid.org/0009-0002-9244-7039
- ORCID[2]: https://orcid.org/0009-0008-4221-3705
- ORCID[3]: https://orcid.org/0009-0001-1811-2903

· Received: Sep. 03, 2026, Revised: Sep. 23, 2026, Accepted: Sep. 26, 2026
· Corresponding Author: Hyun-Joon Chung
Dept. of AI Humanoid R&D Division, Korea Institute of Robotics & Technology Convergence, Korea
Tel.: +54-279-0481, Email: hjchung@kiro.re.kr

Ⅰ. 서 론

초소형 정찰 로봇은 전장, 재난 현장, 산업시설과 같은 위험 지역에서 인간 대신 환경을 탐색하고 상황 정보를 수집하기 위해 활용되고 있다. 특히 야간이나 연기, 먼지, 안개 등으로 가시성이 저하된 환경에서는 일반 RGB 카메라 기반의 시각 인지가 사람 존재 여부를 안정적으로 구분하기 어렵기 때문에, 열영상(Thermal imaging)을 활용한 사람 검출 기능이 필수적이다. 초소형 플랫폼은 크기와 중량, 전력, 연산 자원이 제한적이기 때문에, 이러한 열영상 기반 사람 검출 기능을 임베디드 환경에서 실시간으로 구현하기 위한 경량화·최적화 기술이 함께 요구된다. 열영상 기반 사람 검출에 대한 기존 연구들은 주로 고정형 열화상 카메라나 차량 탑재형 시스템을 대상으로 비교적 여유 있는 연산 자원을 활용하여 높은 검출 성능을 달성하는 데 초점을 맞추어 왔다[1][2]. 예를 들어, FLIR ThermalCAM P10과 같은 중간 이상 해상도의 열화상 센서를 사용해 YOLOv8(You Only Look Once version 8) 계열 모델(nano, small, medium, large, Extra) 검출기를 학습시키고, 정적 장면에서 높은 평균 정밀도(mean Average Precision)를 보고한 연구들이 대표적이다[3]. 한편, 열영상 데이터는 RGB 영상과 달리 해상도가 낮고 윤곽 정보가 약하며, 온도 분포에 따라 배경과 객체의 대비가 크게 달라지는 특성을 지니므로, RGB 기반 사전 학습 모델을 그대로 적용할 경우 성능 저하가 발생하기 쉽다[4]. 이러한 문제를 보완하기 위해 전이학습(Transfer learning)과 도메인 적응(Domain adaptation)을 활용해 RGB에서 학습한 표현을 열영상 도메인으로 전이하는 다양한 방법이 제안되고 있다. 그러나 실제 초소형 정찰 로봇에 탑재 가능한 열영상 사람 검출 시스템을 설계하기 위해서는, 센서와 임베디드 하드웨어의 제약을 고려한 데이터셋 구축과 더불어, 경량화된 검출 모델들이 임베디드 GPU 상에서 어떤 자원 사용량과 추론 속도를 보이는지에 대한 정량적인 분석이 필요하다. Ulmămei et al.[5]은 UAV 열영상 데이터셋을 확장하고 여러 검출 모델을 Jetson 계열 플랫폼에서 비교하여, 모델 구조와 자원 사용량 간 trade-off를 분석하였다. 특히 Jetson Orin NX와 같은 소형 임베디드 GPU 플랫폼은 비교적 강력한 연산 능력을 제공하지만, 군집 로봇 운용 시 복수 센서 처리, 통신, 주행 제어 등이 동시에 수행된다는 점을 감안하면 사람 검출 알고리즘이 사용할 수 있는 연산·메모리 예산은 제한적일 수밖에 없다. 그럼에도 불구하고 초소형 로봇 환경에서 열영상 사람 검출 모델의 정확도, 추론 속도, CPU·GPU 자원 사용량을 체계적으로 비교·분석한 연구는 상대적으로 부족한 실정이다. Tan et al.[6]은 임베디드 환경에서 열 및 적외선 센서를 활용한 사람 검출 시스템을 제안하고, 저속 감시 시나리오를 대상으로 FPGA(Feild Programmable Gate Array) 기반 구현 결과를 보고하였다. 본 연구에서는 초소형 정찰 로봇 탑재를 목표로 FLIR Lepton 3.5 열화상 센서와 PureThermal2 모듈을 활용하여 실제 운용 시나리오를 반영한 열영상 데이터셋 5,000장을 구축하고, 이를 기반으로 다양한 열영상 사람 검출 모델의 임베디드 성능을 분석하고자 한다. 구축된 데이터셋은 낮은 카메라 높이에서 상향 촬영된 시점, 마스크·모자·장갑 등 보호구 착용, 부분 은폐, 실내·실외 및 주변 온도차가 존재하는 환경 등 실제 정찰·경계 임무에서 발생할 수 있는 상황을 포함한다. 또한 Jetson Orin NX 플랫폼을 기준으로 여러 열영상 사람 검출 모델의 검출률, 추론 지연 시간, CPU·GPU 사용량을 비교함으로써, 초소형 정찰 로봇에 적합한 열영상 사람 검출 모델 선택과 임베디드 비전 시스템 설계에 대한 실질적인 기준을 제시한다.

Ⅱ. 열영상 사람 검출 시스템 설계

2.1 열영상 데이터셋 구축

본 연구에서는 초소형 정찰 로봇에 탑재 가능한 열영상 사람 검출 모델을 학습·평가하기 위해, FLIR Lepton 3.5 열화상 센서와 PureThermal2 모듈을 이용하여 열영상 데이터셋을 구축하였다. FLIR Lepton 3.5는 160×120 해상도의 비냉각식 열화상 센서로, 크기(10.5×12.7×7.14 mm)와 무게(0.9g)가 작아 초소형 지상 로봇에 적합하며, PureThermal2 모듈을 통해 USB 인터페이스로 손쉽게 데이터를 수집할 수 있다. 데이터 취득 시스템은 노트북 PC 및 Jetson Orin NX에 연결해 실시간 열영상 스트림을 확인하

면서, 서로 다른 촬영 환경에서 사람 데이터를 캡처할 수 있도록 구성하였다. 데이터셋 구축 시 실제 정찰·경계 상황을 고려하여 다양한 운용 시나리오를 반영하였다. 먼저 초소형 지상 로봇의 낮은 카메라 높이를 모사하기 위해 센서를 지면 가까이 배치한 후, 사람을 아래에서 위로 바라보는 상향 시점 데이터를 수집하였다. 또한 마스크, 모자, 장갑 등 보호구를 착용한 상황, 몸의 일부가 장애물 뒤에 가려진 부분 은폐(Occlusion) 상황, 실내·실외 및 다양한 배경 온도 분포를 포함하여, 열영상에서 사람과 배경 간 온도 대비가 크게 달라지는 조건을 포함하였다. 더불어 학습 안정성을 위해 COCO 및 Roboflow Open 데이터셋과 같은 공개 데이터셋을 함께 활용하고, 실제 환경을 반영한 열영상은 Lepton 3.5 센서로 직접 취득하여 결합하였다. 이러한 구성을 그림 1에 나타낸 것처럼, 왼쪽은 COCO 데이터셋, 가운데는 Roboflow Open 데이터셋, 오른쪽은 Lepton 3.5 센서로 직접 취득한 열영상 데이터의 예시로 정리하였다. 최종적으로 이러한 조건에서 총 5,000장 이상의 이미지를 확보하고, 사람에 대해 bounding box 단위의 라벨링을 수행하였다.

구축된 데이터셋은 학습, 검증, 시험 데이터로 나누어 사용하였다. 학습 및 검증 데이터는 모델 파라미터 학습과 하이퍼파라미터 튜닝에 사용하였고, 시험 데이터는 센서·환경 조건이 학습 데이터와 다소 다른 장면을 포함하도록 구성하여 일반화 성능을 평가하였다. 열영상의 해상도와 시야각 특성을 고려해 단순한 크기 변환과 정규화만을 적용하고, 복잡한 색상 변환이나 스타일 변환은 적용하지 않음으로써 실제 센서 출력과 유사한 조건에서 모델을 평가하도록 설정하였다.

2.2 열영상 데이터의 특성

열영상 데이터는 물체에서 방출되는 적외선을 감지하여 이미지를 생성하므로, 각 픽셀 값이 해당 지점의 온도를 반영한다는 점에서 RGB 영상과 근본적으로 다르다. RGB 영상은 주로 반사되는 가시광선 강도를 표현하는 반면, 열영상은 물체 표면의 절대 온도 또는 온도 차이를 표현하기 때문에, 사람·기계·동물 등 열원과 배경 환경 간의 온도 대비가 중요한 정보로 작용한다. 그림 2는 가시광선 영역과 적외선 영역의 스펙트럼을 개략적으로 나타낸 것으로, 열영상 카메라가 사람 눈으로 볼 수 없는 파장 대역을 감지하여 이미지를 구성한다는 점을 보여준다.

열영상 데이터의 가장 큰 장점은 조명 조건에 독립적이라는 점이다. 적외선 방사는 주변 조명과 상관없이 물체 자체의 온도에 의해 결정되므로, 완전한 어둠이나 낮은 조도 환경에서도 대상의 존재를 안정적으로 감지할 수 있다. 이 특성 덕분에 열영상은 야간 정찰, 침입자 감시, 악천후 상황에서의 상황 인지 등 가시광 카메라가 제대로 작동하기 어려운 응용에서 특히 유용하다. 예를 들어, 연기·먼지·안개 등으로 가시성이 저하된 환경에서도, 열영상은 사람이나 차량과 같이 주변보다 온도가 높은 객체를 비교적 명확하게 분리해 낼 수 있다.

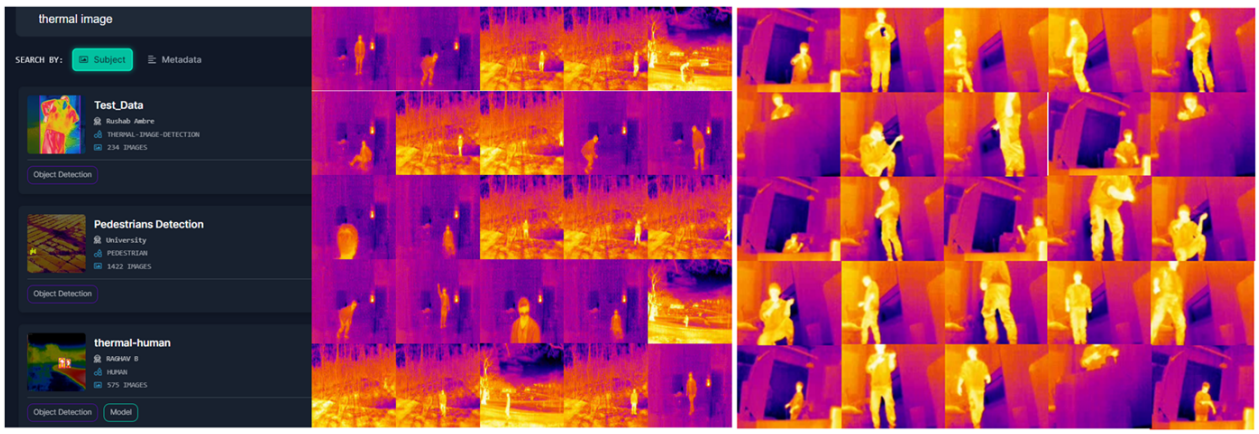


그림 1. 사용된 데이터셋 (왼쪽: COCO 데이터셋, 가운데: Roboflow Open 데이터셋, 오른쪽: Lepton 3.5 센서로 직접 취득한 데이터)

Fig. 1. Dataset used (left) COCO dataset, (center) roboflowopen dataset, (right) captured directly with Lepton3.5 sensor

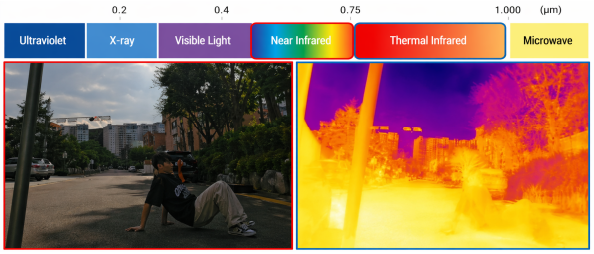


그림 2. 적외선 스펙트럼 및 열영상 데이터 예시
Fig. 2. Infrared spectrum and example of thermal image data

반면 열영상 데이터는 일반적으로 RGB 영상에 비해 해상도가 낮고 세부 텍스처가 부족하다는 한계를 가진다. 이는 센서 구조 및 비용상의 제약으로 인해 열화상 카메라가 상대적으로 적은 픽셀 수를 제공하는 경우가 많기 때문이다. 그 결과, 엣지나 윤곽 정보가 흐릿하게 표현되며, 작은 객체나 서로 인접한 다수 객체를 구분하기 어려운 상황이 자주 발생한다. 또한 낮·밤, 계절, 주변 물질의 열용량과 같은 환경 요소에 따라 배경 온도 분포가 크게 변하기 때문에, 동일한 장면에서도 시간대·기상 조건에 따라 객체와 배경 간 온도 대비가 달라지는 특성이 있다. 이러한 특성은 열영상 기반 사람 검출 모델의 신뢰도(Confidence)가 환경 조건에 따라 민감하게 변동할 수 있음을 의미하며, 본 연구의 실험에서는 야간과 주간 환경에서 이러한 경향을 정량적으로 확인하였다.

2.3 임베디드 비전 환경에서의 검출 모델 구성

열영상 사람 검출 모델은 Jetson Orin NX 임베디드 GPU 플랫폼에서 실시간으로 동작해야 하므로, 정확도와 연산 자원 사용량 간의 균형을 고려한 설계가 필요하다. Jetson Orin NX는 최대 수십 TOPS 수준의 연산 성능과 8~16 GB 메모리를 제공하지만, 실제 로봇 운용 시에는 주행 제어, SLAM (Simultaneous Localization and Mapping) 통신 등 다른 작업과 자원을 공유해야 한다[7][8]. 본 연구에서 사용한 하드웨어·소프트웨어 구성은 표 1과 같이 정리되며, FLIR Lepton 3.5와 PureThermal2를 통해 열영상 데이터를 입력받고, JetPack 5.1 기반 Ubuntu 환경에서 CUDA 11.4 및 OpenCV 4.5.4, PyTorch와 TensorRT를 활용하여 검출 모델을 구동하도록 설계하였다.

검출 모델 후보는 최근 실시간 객체 검출 분야에서 널리 사용되는 단일 단계(One-stage) 검출기 계열을 중심으로 선정하였다. 이들 모델은 입력 이미지를 일정한 격자로 분할하고, 각 격자에서 물체의 존재 여부와 위치를 동시에 추정함으로써, 제안 영역 생성과 분류를 분리하여 수행하는 2단계(Two-stage) 방식에 비해 추론 속도가 빠르다는 장점을 가진다. 각 모델은 동일한 열영상 데이터셋에 대해 사전 학습된 가중치를 기반으로 미세조정(Fine-tuning)하였으며, 입력 해상도, 학습 epoch, 배치 크기 등 학습 설정은 임베디드 배포를 고려하여 통일하였다. 임베디드 구현 환경은 JetPack 기반 Ubuntu OS 상에서 CUDA, cuDNN, TensorRT를 이용해 구성하였다. 학습된 모델은 FP16 및 INT8 등 저정밀도 연산을 지원하는 엔진으로 변환하여 추론에 사용하였으며, 프레임당 추론 시간, CPU 및 GPU 사용률, GPU 메모리 점유량을 각각 측정하였다. 또한 모델 경량화를 위해 필터 가중치 크기를 기준으로 일부 채널을 제거하는 프루닝(Pruning) 기법을 적용하고, 모델 크기 및 파라미터 수 감소 효과와 함께 추론 속도 변화를 관찰하였다. 이러한 설정을 통해, 표 1에 제시된 임베디드 비전 시스템 구성 하에서 각 검출 모델이 실제 로봇 운용 환경에서 어느 정도의 성능과 자원 사용 특성을 보이는지 비교·분석하였다.

표 1. Jetson Orin NX 기반 시스템 구성
Table 1. System configuration on Jetson Orin NX

Category	Item	Details
Hardware	Main board	NVIDIA Jetson Orin NX
	Thermal imaging sensor	FLIR Lepton 3.5
	Interface module	PureThermal2
Software	OS	Ubuntu 20.04
	CUDA	CUDA 11.4
	Jetpack	Jetpack 5.1
	OpenCV	OpenCV 4.5.4
Deep learning / inference	Framework	PyTorch 1.12.0
	Engine	TensorRT 5.1

III. 실험 및 결과

모든 모델은 동일한 열영상 데이터셋을 대상으로

입력 해상도 320×256, batch size 16, 100 epochs 기준으로 학습하였으며, YOLOv8 공식 구현을 기반으로 anchor-free head를 유지한 상태에서 열영상 도메인에 맞춰 fine-tuning을 수행하였다. 학습은 RTX 4090 GPU 환경에서 진행하고, 추론 평가는 Jetson Orin NX에서 FP16 TensorRT 엔진을 사용하여 수행하였다. 실험은 구축된 열영상 시험 데이터셋을 대상으로 각 모델의 검출률과 임베디드 자원 사용량을 분석하는 방식으로 진행하였다. 검출 성능 평가는 평균 정밀도(mAP)를 기준으로 수행하였으며, 임계값은 일반적인 객체 검출 연구에서 사용되는 IoU 0.5 기준을 적용하였다. 임베디드 성능 평가는 Jetson Orin NX에서 1장당 평균 추론 시간(ms), GPU 메모리 사용량, CPU·GPU 사용률을 지표로 삼았으며, 그 결과는 표 2에 정리하였다. 표 2에서 보듯이 nano, small, medium, large, extra large 크기의 검출 모델에 대해 자원 사용량과 검출 성능을 한눈에 비교할 수 있다.

표 2. Jetson Orin NX 기반 열영상 사람 검출 모델별 자원 사용량 및 성능 비교
Table 2. Resource usage and performance of thermal human detection models on Jetson Orin NX

Model version	CPU usage(M)	GPU usage(G)	Accuracy (%)	Speed (ms)
nano	828	1.2	82.0	25
small	828	1.3	82.7	51
medium	830	1.4	82.7	98
large	831	1.5	82.3	146
extra	829	1.6	82.6	221

실험 결과, 가장 경량인 nano 모델은 최소 수준의 CPU·GPU 자원을 사용하면서도 프레임당 25 ms의 매우 빠른 추론 속도를 보였으나, 비교적 낮은 검출률을 보여 일부 열악한 환경에서 사람을 놓치는 사례가 확인되었다. 반면 small 및 medium 모델은 검출률이 82.7%로 향상되면서도 추론 시간(51 ms, 98 ms)과 GPU 메모리 사용량(각각 약 1.3 GB, 1.4 GB)이 여전히 임베디드 환경에서 수용 가능한 범위에 머물러, 실제 초소형 정찰 로봇에 탑재하기에 적합한 후보로 나타났다. 가장 대형인 large 및 extra large 모델의 경우 검출률은 소폭 향상되었으나, GPU 메모리 점유와 추론 시간이 각각 146 ms, 221 ms 수준으로 크게 증가하여 실시간성이 요구되는 임무에는 적합하지 않은 것으로 판단되었다.

정량적인 지표 외에도, 실제 정찰·경계 시나리오에서 열영상 기반 검출이 가지는 장점을 확인하기 위해 야간 환경과 다양한 자세에 대한 정성적 평가를 수행하였다. 그림 3은 야간 환경에서 가시광 카메라(좌)와 열영상 카메라(우)를 각각 사용하여 사람 검출을 수행한 결과를 비교한 것이다. 가시광 카메라 영상에서는 조도가 낮고 배경 조명으로 인해 사람과 배경의 경계가 흐려져 검출이 불안정한 반면, 열영상에서는 사람과 배경 간 온도 차이가 뚜렷하게 나타나며, 어두운 환경에서도 사람 위치를 안정적으로 탐지할 수 있음을 확인할 수 있다.

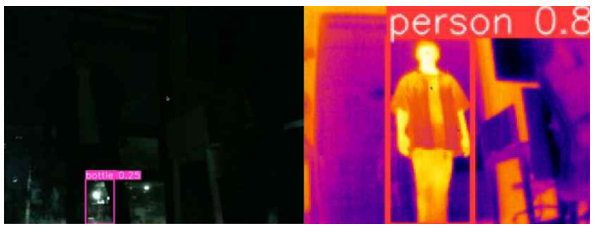


그림 3. 야간 환경에서 가시광 카메라(좌)와 열영상 카메라(우)를 이용한 사람 검출 결과 비교
Fig. 3. Comparison of human detection in a night-time scene using a visible camera (left) and a thermal camera (right)

그림 4는 야간(좌)과 주간(우) 환경에서 동일 위치의 사람을 대상으로 열영상 기반 검출을 수행한 결과를 나타낸다. 두 경우 모두 사람에 대한 검출 자체는 성공하였으나, 주간 환경에서는 배경의 온도가 상승하여 주변 물체와 사람 사이의 온도 대비가 감소함에 따라, 사람 영역의 신뢰도(Confidence) 값이 야간에 비해 약 0.5 수준으로 낮게 나타났다. 이는 낮 시간대에는 지면과 주변 구조물에 태양 복사열이 축적되면서 전체적으로 열 신호가 포화되고, 사람의 체온과 배경 온도 차이가 줄어들어 모델이 사람을 뚜렷한 고온 영역으로 인식하기 어려워지는 영향으로 해석할 수 있다. 반면 야간에는 배경이 빠르게 냉각되어 사람과 주변 환경 간 온도 대비가 커지므로, 동일한 모델 설정 하에서도 더 높은 신뢰도(Confidence)로 사람을 검출하는 경향을 보였다. 이러한 결과는 열영상 기반 사람 검출에서 야간·주간에 따라 배경 온도 분포가 달라질 경우, 임계값 설정과 학습 데이터 구성에 차별화된 고려가 필요함을 시사한다.

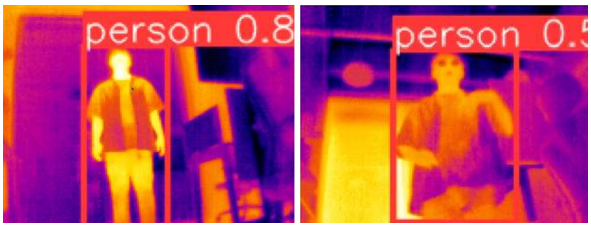


그림 4. 야간/주간 환경에서 열영상 기반 사람 검출 결과 (좌: 야간, 우: 주간)
Fig. 4. Thermal human detection results in night-time and day-time scenes (left: night, right: day)

프루닝을 적용한 모델은 파라미터 수와 모델 파일 크기가 감소하는 효과는 있었으나, GPU 상에서 여전히 동일한 크기의 텐서 연산이 수행되기 때문에 프레임당 추론 시간은 기대만큼 줄어들지 않았다[9]. 그림 5는 신경망에서 프루닝을 적용할 때, 중요도가 낮은 연결과 뉴런을 제거하여 네트워크 구조를 간소화하는 개념을 나타낸 것이다. 이러한 방법을 통해 모델의 파라미터 수와 저장 용량은 줄어들지만, 구조적으로는 동일한 크기의 특성 맵과 텐서가 GPU에서 연산되기 때문에 실제 합성곱 연산량과 메모리 접근 패턴에는 큰 변화가 없다. 이는 이미 경량화된 모델 구조에서는 단순 파라미터 제거만으로는 연산량 감소가 충분히 이루어지지 않음을 시사하며, 구조 자체를 변경하는 네트워크 설계나 하드웨어 수준에서 희소성을 활용할 수 있는 보다 적극적인 최적화 기법이 필요함을 의미한다.

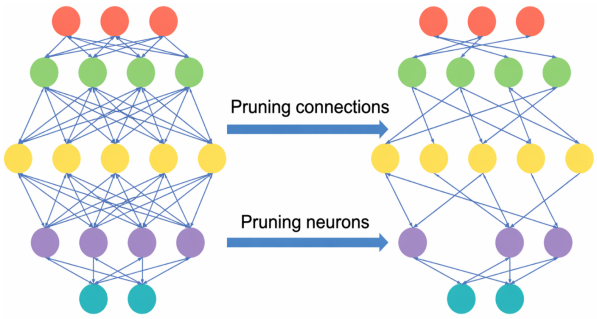


그림 5. 신경망에서 연결 및 뉴런 제거를 통한 프루닝 개념도
Fig. 5. Example of neural network pruning by removing connections and neurons

Ⅳ. 결론 및 향후 과제

본 연구에서는 초소형 정찰 로봇에 탑재 가능한 열영상 기반 사람 검출 시스템을 구현하기 위해, FLIR Lepton 3.5 및 PureThermal2 모듈을 활용한 열영상 데이터셋을 구축하고 Jetson Orin NX 임베디드 플랫폼 상에서 여러 검출 모델의 성능을 비교·분석하였다. 실제 정찰·경계 시나리오를 고려하여 카메라 높이, 보호구 착용, 부분 은폐, 실내·실외 및 배경 온도 조건을 다양화한 5,000장 이상의 열영상 데이터를 확보함으로써, 초소형 로봇 운용 환경에 특화된 학습·평가 기반을 마련하였다. 또한 nano, small, medium, large, extra large 크기의 검출 모델에 대해 열영상 사람 검출률, 추론 속도, CPU·GPU 자원 사용량을 정량적으로 비교하고, 야간·주간 및 저조도 환경에서의 정성적 검출 결과를 함께 분석하였다.

실험 결과, 가장 경량인 nano 모델은 프레임당 25 ms 수준의 매우 빠른 추론 속도와 낮은 자원 사용량을 보였으나, 일부 열악한 환경에서 검출률이 낮아 정찰·경계 임무에 단독으로 사용하기에는 한계가 있었다. 반면 small 및 medium 모델은 82.7%의 검출률과 임베디드 환경에서 수용 가능한 추론 시간·GPU 메모리 사용량을 동시에 달성하여, 초소형 정찰 로봇용 열영상 사람 검출 모델로 가장 균형 잡힌 성능을 제공하는 것으로 확인되었다. large 및 extra large 모델은 검출률이 소폭 향상되었지만, 146 ms 및 221 ms 수준의 추론 지연과 높은 자원 사용량으로 인해 실시간성이 요구되는 임무에는 적합하지 않은 것으로 나타났다. 야간·주간 환경 비교 결과에서는, 주간에 배경 온도가 상승하면서 사람과 배경 간 온도 대비가 감소하여 사람 영역의 신뢰도(Confidence)가 야간 대비 낮게 나타나는 현상이 관찰되었으며, 이를 통해 열영상 기반 정찰 시스템에서도 시간대별 배경 온도 분포를 고려한 임계값 설정과 데이터 구성의 중요성을 확인하였다.

프루닝 실험에서는 파라미터 수와 모델 파일 크기는 감소했으나, GPU 상에서 처리되는 텐서 크기와 연산 그래프 구조가 유지됨에 따라 프레임당 추론 시간은 기대만큼 개선되지 않았다. 이는 이미 경량화된 검출기 구조에서는 단순 가중치 제거만으로는 연산량과 메모리 접근 패턴을 충분히 줄이기 어렵다는 점을 보여주며, 구조적 변경을 수반하는 경량 네트워크 설계나 하드웨어 친화적 최적화 기법이 병행되어야 함을 시사한다. 예를 들어, 임베디드

플랫폼에 특화된 경량 백본 네트워크, 채널·레이어 수준의 재구성, TensorRT 기반 연산 그래프 최적화 및 커널 융합 등을 적용할 경우 추가적인 추론 속도 향상을 기대할 수 있다.

향후 연구에서는 열영상 도메인의 특성을 반영한 전용 네트워크 구조 설계와 더불어, RGB와 열영상 간 도메인 갭을 완화하기 위한 도메인 적응 기법을 결합하여 다양한 환경에서의 검출 성능을 향상시키고자 한다. 또한 군집 정찰 로봇 운용을 고려하여, 다수 로봇에서 동시에 열영상 사람 검출, SLAM, 통신 기능을 수행할 때의 작업 스케줄링 및 자원 관리 전략을 포함하는 시스템 수준 최적화에 대한 연구가 필요하다. 마지막으로, 실제 실내·실외 시험 환경에서 장시간 운용 데이터를 수집·분석하여, 본 논문에서 제시한 모델 선택 기준과 임베디드 구성 방안이 다양한 정찰·경계 시나리오에서 얼마나 안정적으로 작동하는지 검증하는 작업이 요구된다.

References

[1] S.-M. Yang and J.-K. Choe, "A study on improvement of human sensing system reliability using thermal image", J. Korea Inst. Inf. Commun. Eng., Vol. 18, No. 1, pp. 73-81, Jan. 2020. https://doi.org/10.14801/jkiit.2020.18.1.35.

[2] J.-H. Kim and J.-M. Park, "Intrusion detection and monitoring system using webcam and infrared thermal camera module", J. Internet Comput. Serv., Vol. 25, No. 5, pp. 77-86, Oct. 2024. https://doi.org/10.12673/jant.2024.28.5.754.

[3] M. Ivašić-Kos, M. Krišto, and M. Pobar, "Human detection in thermal imaging using YOLO", Proc. 2019 5th Int. Conf. on Computer and Technology Applications, New York, NY, USA, pp. 45-49, Apr. 2019. https://doi.org/10.1145/3323933.3324076.

[4] H.-S. Lee, "A study on infrared thermal imaging measurement in high-temperature environments", J. Korea Soc. Mech. Eng., Vol. 48, No. 4, pp. 325-333, Apr. 2024.

[5] A. A. Ulmămei, T. D'Adamo, C. E. Vasile, and R. Hobincu, "Human detection in UAV thermal imagery: Dataset extension and comparative evaluation on embedded platforms", J. Imaging, Vol. 11, No. 12, p. 436, Dec. 2025. https://doi.org/10.3390/jimaging11120436.

[6] T. F. Tan, S. S. Teoh, J. E. Fow, and K. S. Yen, "Embedded human detection system based on thermal and infrared sensors for anti-poaching application", Proc. 2016 IEEE Conf. on Systems, Process and Control (ICSPC), Malacca, Malaysia, pp. 37-42, Dec. 2016. https://doi.org/10.1109/SPC.2016.7920700.

[7] Y.-H. Kim, D.-Y. Lee, and J.-H. Kwon, "Conversion and lightweight optimization of object detection models for efficient operation in embedded environments", J. Korea Inst. Inf. Commun. Eng., Vol. 27, No. 9, pp. 1143-1152, Sep. 2023.

[8] H.-J. Shin and S.-G. Lee, "Trends of mobile and embedded object and scene recognition technologies", ETRI J., Vol. 34, No. 6, pp. 133-144, Dec. 2012. https://doi.org/10.22648/ETRI.2019.J.340612.

[9] S. Belhadi, R. Rollet, A. Genty, and V. Gripon, "LightPrune: Latency-aware structured pruning for efficient deep neural networks on embedded systems", Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW), Paris, France, pp. 123-132, Oct. 2025. https://doi.org/10.1109/ICCVW69036.2025.00179.

저자소개

강 전 성 (Jeon-Seong Kang)



2018년 2월 : 동국대학교 전자전기공학과(공학석사)
2018년 1월 ~ 2021년 5월 : ㈜ 렉스젠 AI 전문연구요원
2021년 9월 ~ 현재 : 한국로봇융합연구원 AI·휴머노이드연구본부 주임연구원
관심분야 : Computer Vision, VLA, Robotics, Embedded System

박 범 준 (Beom-Joon Park)



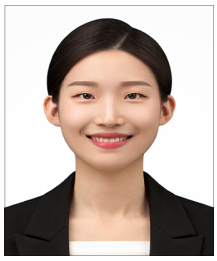
2017년 2월 : 인하대학교 전기공학과(학사)
2019년 2월 : 연세대학교 전기전자공학과(공학석사)
2017년 3월 ~ 2020년 12월 : 한국과학기술연구원 연구원
2022년 8월 ~ 현재 : 한국로봇융합연구원 AI·휴머노이드연구본부 주임연구원
관심분야 : Reinforcement Learning, Swarm Intelligence, Motion Planning, Planetary Robotics, Legged Robotics, Field Robotics

윤 준 원 (Junwon Yoon)



2020년 2월 : 고려대학교 기계공학과(학사)
2022년 8월 : 고려대학교 기계공학과(공학석사)
2023년 7월 ~ 현재 : 한국로봇융합연구원 AI·휴머노이드연구본부 주임연구원
관심분야 : Robotics, Biomechanics, Generative Model, Vision-Language-Action Model, Few-shot Learning, Data Augmentation

송 하 윤 (Ha-Yoon Song)



2020년 2월 : 가톨릭대학교 심리학과(학사)
2023년 2월 : 고려대학교 뇌공학과(석사)
2023년 11월 ~ 현재 : 한국로봇융합연구원 AI·휴머노이드연구본부 주임연구원
관심분야 : Human-Robot Interaction, Brain-Computer Interface, Vision-Language-Action Model, Multimodal Learning

정 현 준 (Hyun-Joon Chung)



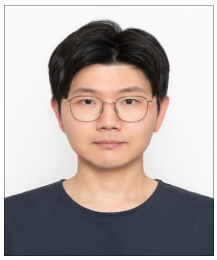
2005년 5월 : University of Iowa 기계공학과(공학석사)
2009년 12월 : University of Iowa 기계공학과(공학박사)
2015년 11월 ~ 2021년 1월 : 한국로봇융합연구원 선임연구원
2021년 2월 ~ 현재 : 한국로봇융합연구원 AI·휴머노이드연구본부 책임연구원, 본부장
관심분야 : Dynamics and Control, Optimization Algorithms, Computational Decision Making, Robotics, Modeling and Simulation

서 덕 현 (Duck-Hyun Suh)



2020년 8월 : 광운대학교 로봇학부 정보제어전공(학사)
2022년 8월 : 고려대학교 전기전자공학부 제어, 로봇, 시스템 전공(석사)
2022년 10월 ~ 현재 : LIG Defense & Aerospace 선임연구원
관심분야 : LLM, VLM, VLA, Mission Planning, Multi-modal learning

윤 태 준 (Taejoon Yoon)



2017년 2월 : 경희대학교 기계공학부(학사)
2019년 2월 : 경희대학교 기계공학부(석사)
2019년 6월 ~ 현재 : LIG Defense & Aerospace 선임연구원
관심분야 : LLM, VLM, VLA, Robotics SW Architecture, Mission Planning

김 주 영 (Jooyoung Kim)



2007년 2월 : 이화여자대학교 컴퓨터공학과(학사)
2009년 1월 : 이화여자대학교 컴퓨터정보통신학과(석사)
2009년 1월 ~ 현재 : LIG Defense & Aerospace 수석연구원
관심분야 : Autonomous Control, Swarm Control, LLM