

군 로봇의 장소 분류 정확도 향상을 위한 적외선 이미지 데이터 결합 학습 방법 연구

A Study on the Training Methodology of Combining Infrared Image Data for Improving Place Classification Accuracy of Military Robots

최 동 규¹ · 도 승 원¹ · 이 창 은[†]

Donggyu Choi¹, Seungwon Do¹, Chang-eun Lee[†]

Abstract: The military is facing a continuous decrease in personnel, and in order to cope with potential accidents and challenges in operations, efforts are being made to reduce the direct involvement of personnel by utilizing the latest technologies. Recently, the use of various sensors related to Manned-Unmanned Teaming and artificial intelligence technologies has gained attention, emphasizing the need for flexible utilization methods. In this paper, we propose four dataset construction methods that can be used for effective training of robots that can be deployed in military operations, utilizing not only RGB image data but also data acquired from IR image sensors. Since there is no publicly available dataset that combines RGB and IR image data, we directly acquired the dataset within buildings. The input values were constructed by combining RGB and IR image sensor data, taking into account the field of view, resolution, and channel values of both sensors. We compared the proposed method with conventional RGB image data classification training using the same learning model. By employing the proposed image data fusion method, we observed improved stability in training loss and approximately 3% higher accuracy.

Keywords: Infrared Image, Military Robots, Classification, Dataset

1. 서 론

현재 국내 인구수 문제로 인한 군 인력 문제가 지속 제기되고 있으며, 이러한 문제를 해결하기 위해 다양한 구조 개편 및 신기술 도입 등의 정책을 시행하고 있다. 더불어, 작전 수행에 있어 물리적인 인력이 투입됨과 동시에 작전 위험성에 따라 심각한 문제로 발전할 수 있다. 군 작전 사고로 인한 인원 부채 시 기존 인력의 업무가 가중으로 인하여 실수 가능성이 증가

하고 상황 제어와 대응 능력이 저하될 수 있다. 또한, 작전 수행 시 지원 및 후속 조치도 중요한 요소이며 정비 및 보급과 구조와 같은 행위가 충분히 이루어지지 어렵다¹⁻³⁾.

군의 작전은 적침 및 위협에 따른 신속한 기동 타격을 위하여 주야간을 가리지 않고 수행된다. 하지만, 주간에는 사람이 육안으로 적을 식별하고 문제상황에 관해 대응할 수 있으나 야간의 경우 시각 정보가 매우 부족하여 잠재적인 위협요소를 놓치거나 적의 위치를 파악하기 쉽지 않다. 이러한 어두운 환경은 주간뿐만 아니라 빛이 부족한 실내, 밀폐환경 등의 환경을 가진 작전 상황을 포함하게 된다. 또한, 피로와 수면 부족과 같은 인체의 신체적 한계에 따라 즉각적인 동작과 유연한 대처가 어려울 수 있다. 전반적으로 어두운 작전환경은 위험성이 높다. 적의 공격이 예상하기 어려우며, 미세한 감각에도 혼란이 증대하게 된다. 추가로 복잡한 지형과 장애물까지 더해진다면 안전한 이동과 탐색이 어렵다^{4,5)}.

Received : May. 31. 2023; Revised : Jun. 26. 2023; Accepted : Jun. 28. 2023

※ This work was supported by Korea Research Institute for Defense Technology planning and advancement (KRIT) grant funded by Korea government DAPA (Defense Acquisition Program Administration) (No. KRIT-CT-22-006-002, Development of the situation/environment recognition technology for micro-swarm robot)

1. Researcher, ETRI, Daejeon, Korea (dgchoi, seungwon.do@etri.re.kr)

† Principal Researcher, Corresponding author: Defense ICT Convergence Research Section, ETRI, Daejeon, Korea (celee@etri.re.kr)

미국에서는 군에 다양한 인공지능 기술을 도입하여 인력에서 오는 문제와 편의를 해결하기 위하여 시스템을 개발하고 있다. 2017년에 미 국방부는 Algorithmic Warfare Cross-Functional Team을 시작으로 2019년 합동 AI센터(JAIC)를 신설하였다. 대표 분야로는 감시정찰, 사이버 작전, 정보작전, 지휘통제, 반자율/자율 차량, 자율살상무기로 지정하고 각 분야에 맞는 원천기술을 개발하고 있다. 또한, 미 국방부 산하의 DARPA에서는 2014년부터 자동 방어 시스템을 개발하기 위하여 Cyber Grand Challenge를 개최하고 있으며, 2016년 AI 기반의 사이버 도구를 활용한 팀이 공격과 방어를 동시에 수행하여 우승함에 따라 그 가능성을 확인하였다⁶⁾.

국내에서는 이러한 난제를 해결하기 위해 TOD나 고성능 CCTV와 같은 장비가 도입되어 활용되고 있으며 넓은 영역의 감시 정찰을 수행하여 일부 인력을 대체하고 있다⁷⁾. 하지만, 실제 작전 및 실내 상황에서 활용되지 않으며 아직 사람의 개입이 많고, 고정된 위치를 감시하며 이미 설치된 위치를 알고 작전을 수행한다. 해당 시스템들은 크기가 크고 수동 조작으로 수행되기에 일반적 상황에서 주로 운용되므로 실제 작전마다 필요한 형태의 기기와 자동화 기술을 적용하게 된다면, 자율적인 장소 인식과 더불어 많은 문제의 해결이 필요하다.

본 논문에서는 앞선 문제점들을 해결하기 위하여 군 로봇 제작에 필요한 인공지능의 시각적 인지 기능의 데이터 학습 방법을 연구하였다. 결과로는 로봇의 이동에 따라 변화는 장소를 분류할 수 있도록 인공지능 모델인 Vision Transformer를 경량 설계한 자체 모델을 활용하였다. 학습에 필요한 데이터의 정보량을 증대하기 위해 기존 RGB 이미지와 더불어 IR 이미지를 새로운 채널로 구성하여 활용하게 되며, 동일 장소의 어두운 환경 데이터도 학습에 사용된다.

2. 관련 연구

본 논문에서는 제안 방법을 진행하기에 앞서 학습에 주로 사용되는 공공 데이터셋과 필요한 데이터셋의 특징과 장소 분류 내용, RGB 및 IR 이미지 데이터의 딥러닝 활용과 관련한 기존 연구의 내용을 확인하였다.

2.1 Place365 dataset

일반적으로 학습 테스트 시에 공공 데이터셋을 많이 활용하게 된다. 로봇의 환경 및 장소 분류에 활용하기 위해서는 그러한 분류 작업과 관련한 데이터가 필요하다. 장소 분류와 관련한 대표적인 이미지 데이터셋으로 Place365 데이터셋이 있다.

[Fig. 1]은 Place365 데이터셋에 포함되어 있는 예시 이미지를 나타낸다. Place365 데이터셋은 다양한 장소와 환경을 포함



[Fig. 1] Example images of a Place365 dataset^[8]

하고 있는 대규모 이미지로 구성되어 있다. 이 데이터셋은 주로 장소 분류 작업에 활용되며, 365개의 장소 카테고리 및 약 1.8백만 개의 이미지를 포함하고 있다. 각 이미지는 하나의 장소 카테고리로 Labeling되어 있어 이미지 분류 작업에 적합하다. 또한, 다양한 이미지 데이터의 시각적 특성으로 인하여 특징을 반영해 군 작전 로봇의 장소 분류에도 활용될 수 있다⁸⁾.

하지만, Place365 데이터셋은 해외의 데이터이며 한국의 환경을 학습할 건물이나 장소는 상대적으로 적다. 또한, 대부분 주간 촬영된 이미지를 포함하고 있기에 주간 환경에 대한 분류 작업에 적합하다. 군 작전은 주야간을 가리지 않고 수행하는 상황을 고려할 때, 야간 작전 환경 분류에는 부족한 부분이 있을 수 있다. 군 작전 로봇의 다양한 야간 환경 인지 정확성을 높이기 위해서는 군에서 활용하는 IR로 감지된 이미지를 활용하는 것이 필요하다.

2.2 Infrared image

Infrared (IR) image는 일반적으로 적외선이 감지되어 나타내어진 이미지를 칭하며, 현재 2가지의 활용 방식이 있다. 열 신호를 감지하는 센싱 방법과 적외선 영역의 전자기파를 감지하는 센싱 방법이 있다.

[Fig. 2]은 IR 이미지의 2가지 방식으로 촬영된 결과를 나타낸다. 좌측에 해당하는 적외선을 감지하는 방식은 적외선 영역의 전자기파를 감지하는 방법으로 보조 장치인 IR laser projector를 이용하여 적외선 빛을 방출해 주변 환경을 조명하고, 객체에서 반사되어 돌아오는 빛을 감지하여 이미지를 취득한다. 이



[Fig. 2] Laser IR Image and Thermal Sensing IR Image^[9]

러한 정보를 활용해 객체의 표면 특성 및 깊이 측정 등에 활용이 가능하다^{9,10}.

주로 군 작전에서 사용되는 것은 [Fig. 2]의 우측 이미지인 열 신호를 감지하는 방법이다. 이는 야간에도 사람이나 차량, 건물 등의 열패턴을 확인하여 식별이 가능하고, 단순한 물체 감지뿐 아니라, 숨은 적이나 야간환경 인식이 가능하다. 두 방법 모두 일반적인 사람의 시각으로 확인할 수 없는 데이터 정보를 취득하게 되며 야간에서 활용 가능하다는 특징이 있다. 일반적인 영상과 같은 데이터와 다르지만 같은 곳을 촬영하는 영상에서 추가적인 정보를 제공할 수 있다.

2.3 Research on RGB and IR Image Fusion Learning

일반적으로 인공지능 모델의 학습을 수행하는 것은 RGB 이미지 데이터를 활용하고 있으며 IR 이미지 데이터는 Public data가 많지 않고, 특수 센서 사용이 필요하므로 데이터 확보나 취득이 어렵다. 하지만, 학습 모델의 정확도 향상을 위하여 일부 사용되어 왔다.

2019 IEEE ITSC에서 발표된 ‘Deep Convolutional Neural Network-based Fusion of RGB and IR Images in Marine Environment’ 논문은 해양 환경에서 활용하기 위하여 RGB와 IR 이미지 데이터를 개별적으로 RetinaNet 모델을 사용하여 특징점을 추출한 후 결합하는 다중 프레임워크를 제안한다. 해당 논문에서는 Parameter 수를 언급하고 있지 않아 개별 수행에 따라 약 1.5배 이상 연산 가중이 일어날 것으로 보인다. 하지만, RGB와 IR의 단일 데이터 활용 검출이나, 학습 모델의 중간 융합 검출 방법에 비해 개별적 모델 연산 이후 결합을 통해 여객선을 감지하는 것에 있어서 약 3.7% 정확도가 향상되는 것을 확인하였다¹¹.

2022년에 Journal of Field Robotics에 게재된 ‘Deep learning with RGB and thermal images onboard a drone for monitoring operations’ 논문은 드론에 탑재하기 위하여 RGB 및 IR 이미지를 사용한 딥러닝 모델 및 방법을 제안한다. 해당 논문에서는 일본에서의 지진 재난 피해로 인한 구호 제공 방법을 모색하기 위하여 빠른 모니터링 및 순찰을 위해 드론을 활용하고 기기와 접목할 딥러닝 기술을 서술하고 있다. 방법으로는 RGB와 IR 이미지 데이터 각각 딥러닝을 일부 수행한 이후 특징점을 결합하는 것과 데이터를 결합하여 4채널 이미지로 구성하여 학습한 결과를 나타낸다. 모델로는 YOLOv3를 사용하였으며, 4채널 이미지 데이터로 구성하여 학습한 결과가 IR 영상만을 활용한 것보다 정확도가 약 5.8% 향상되는 것을 확인하였다¹².

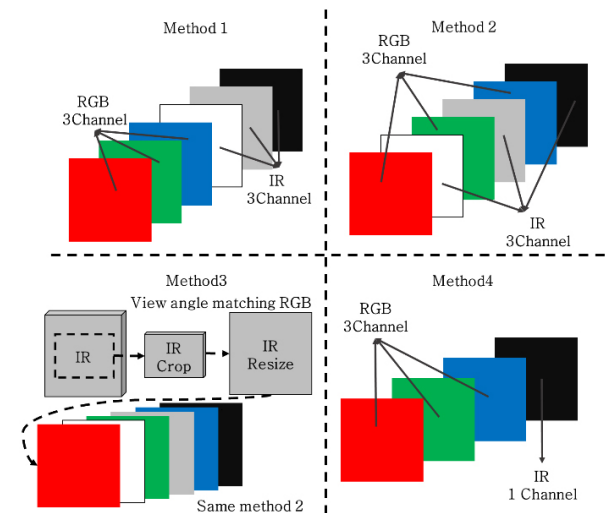
대체로 RGB와 IR을 단일로 사용하거나, 결합하기 위해서는 각각의 데이터를 개별적으로 다중 프레임워크로 학습하는 방법을 사용하여 인식 및 인지에 필요한 작업을 수행하도록

연구되고 있다. 해당 연구들은 객체 인식 수행에서 효과를 보이는 것은 확인되나, 분류나 장소 인식과 관련한 내용은 아니며 군 로봇의 경우 연산이 빠른 모델을 통해 동적으로 실내를 포함하여 장소를 인식할 필요가 있다.

3. 설계 방법

본 논문에서 수행한 방법으로는 4개로, 사용되는 데이터가 이미지인 특성에 따라 채널 변형을 활용하여 설계되었다. 각각의 데이터를 개별적으로 학습을 수행하기 위하여 모델을 구성하면 연산에 필요한 Parameter 수가 증가하여 하드웨어 리소스를 많이 사용하기에 가장 직관적인 해당 방법을 활용하였다. 사용된 데이터는 일반적인 RGB 데이터와 동일 장소에서 취득된 적외선 감지 방식의 IR 데이터를 활용하였다. IR 데이터는 RGB 데이터의 표시방식인 0~255 값과 달리 변환된 깊이 값이 확인되나, 이를 Normalization하여 해당 영역값으로 변환하여 결합에 활용되었다.

[Fig. 3]의 경우 본 논문에서 제안하는 4가지의 학습 수행 전 데이터 변형의 방법으로 좌측 상단부터 오른쪽 하단까지 1~4번의 방법을 나타내고 있다. 일반적인 데이터의 사용은 3채널의 RGB 데이터로 구성되며, IR 데이터 추가로 인하여 최소 4~6개의 채널로 구성되어 활용된다. 1번의 방법은 RGB와 IR 데이터를 동일한 3차원으로 동일하게 맞추어 총 6차원의 데이터로 결합하는 것을 의미한다. 2번의 방법은 1번의 방법에서 RGB와 IR의 채널을 분리하여 교차시키는 방법으로 RGB와 IR 데이터를 순서별로 1채널씩 결합하는 것을 의미한다. 3번의 방법은 RGB와 IR 이미지 센서에서 취득되는 다른 시야각을 같은 크기로 맞추기 위하여 Crop과 Resize를 수행하여 6차



[Fig. 3] Proposed methods



[Fig. 4] Example images of the dataset

원의 데이터로 변환하여 채널 데이터를 섞어 결합하는 것을 의미한다. 4번의 방법은 가장 간단하게 IR 데이터를 1채널로 하여 RGB와 함께 총 4차원 데이터로 결합하는 것을 의미한다.

[Fig. 4]는 학습에 활용된 데이터셋의 Class 중 Corridor에 해당하는 데이터의 예시를 나타낸다. [Fig. 2] 상단에 위치한 이미지는 RGB 데이터이며, 하단은 IR 데이터를 나타낸다. 데이터의 특징으로는 원본 이미지를 비교하였을 때 각 이미지 센서에 탑재된 렌즈가 달라 화각이 다른 것을 확인할 수 있다. 원본의 경우 가로가 긴 화각을 가지는 이미지로, 최소한의 변형을 통하여 학습 연산에 용이하도록 가로와 세로를 동일한 크기로 변형해야 한다. 그에 따라 데이터는 Crop과 Resize를 수행해 비교하였으며, [Fig. 4]의 왼쪽부터 원본, Crop, Resize를 수행한 데이터를 나타낸다. 데이터를 사용하기 전 가장 유사하게 결합할 수 있는 형태로는 RGB의 경우 Resize와 IR의 경우 Crop의 형태가 육안상으로 가장 적합한 것으로 확인되었다.

$$H = - \sum_{i=0}^{255} p_i \log_2 p_i \quad (1)$$

식 (1)은 이미지의 로컬에 1개의 픽셀에 해당하는 정보량은, 복잡성을 나타내는 엔트로피 계산식을 나타낸다^[13]. 식 (1)에서 p_i 는 이미지의 정규화된 히스토그램에서 얻어지는 Gray Scale에 관한 확률 값을 의미한다. [Fig. 2]에서 원본 이미지의 엔트로피 계산을 수행했을 때 RGB에서는 6.17694831, IR 이미지에서는 6.5232583 값을 결합한 이미지에서는 6.80524607의 값을 보여준다. 값이 높을수록 더 많은 정보를 담고 있다는 것을 의미하고 일반적으로 색상 데이터가 더 많은 특징점을 지니고 있으나, 단순히 기존 데이터에 추가 데이터를 활용하여도 얻는 정보량이 수치상으로 더 많아지는 것을 알 수 있다.

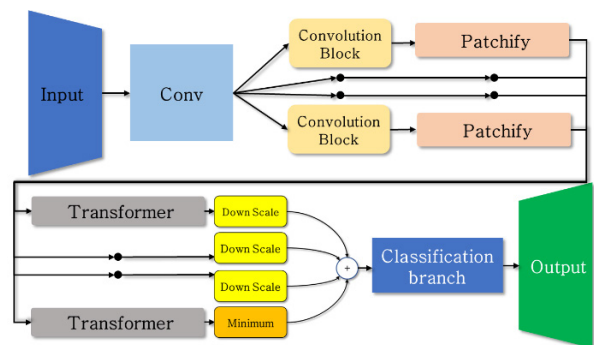
[Table 1]는 학습에 활용된 데이터셋의 사양을 나타낸다. 자체 수집한 데이터셋이 사용되었으며, 기본적으로 RGB 데이터를 사용하나 동일한 형태의 IR 데이터를 활용 시 양이 2배로 증가한다. 일반적으로 건물 안에서 촬영된 이미지 데이터이며, 분류 모델 학습에 활용할 수 있도록 Labeling 작업을 수행하였다.

[Table 1] Specifications of the utilized dataset

Resolution	640*480
Data size	9,338
	Combining IR images results in a twofold increase in quantity
	20% of the total dataset // 1,868
Class	8 (Corridor, Elevator, Parking, Robby_stair, seminar_room, seminar_room_door, stair, table_tennis_room)
Purpose	Classification
Place	Building (Inside)

4. 결 과

본 논문에서 진행한 학습 데이터 변형에 따른 인공지능 모델은 직접 설계한 것으로 [Fig. 5]과 같은 구조를 지닌다. 입력 데이터로 제안하는 방법들을 활용하게 되며, 각 방법마다 채널 크기가 다르기 때문에 이를 활용할 수 있는 크기로 변형하는 Convolution layer를 통과한다. 이후 기존 Vision Transformer의 경량화를 위해 사용된 Convolution block을 거쳐 데이터를 Patch로 쪼갬 후 Transformer 과정을 수행한다. Convolution block은 병렬구조로 설계되어 크기가 다른 4개의 데이터로 변형하게 된다. Transformer까지 수행한 이후 데이터의 크기가 다르기에 맞추기 위해 제일 크기가 작은 Transformer 수행 결과의 크기를 확인하고 나머지 3개의 Transformer 결과를 Down Scale Convolution을 거쳐 Classification branch를 수행하여 분류 결과를 보여준다. 학습 검증에는 StratifiedKFold를 이용한 교차 검증 방법을 사용하여 불균형적인 Class 문제 해결과 Validation에 사용되는 데이터를 epoch마다 변경할 수 있도록 활용하였다. 모델의 경량화는 이후 로봇 플랫폼에 사용할 수 있도록 고려하여 진행하고 있으며, 기존 ViT와 크게 다른 점은 Patch 개수를 절반으로 하는 것과 Transformer를 4개의 데이터 사이즈를 활용하여 Parameter 수를 매우 줄였다.



[Fig. 5] Custom model designed for testing the proposed method

[Table 2] Hardware specifications

OS	Windows 10
CPU	Intel core i7-13700 K 3.40 GHz
GPU	NVIDIA RTX 4080 16 GB
RAM	32 GB DDR5

[Table 3] Results of the proposed methods

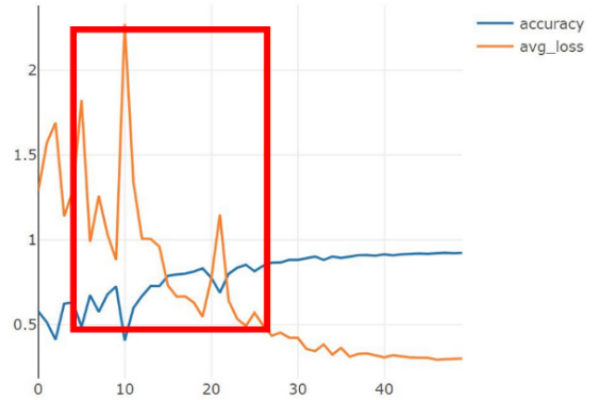
Data type	RGB + IR Image (Size:Basic)		
Model	Vision Transformer (Custom)		
Parameter	Learning late:0.001 / Epoch:50 / Batch size:20		
Data size	18,676 (RGB + IR / RGB:9,338)		
Method	Image size	Input	Accuracy
	256 x 256 (Basic)	RGB(3ch)	92.2912
		RGB(3ch) + IR(1ch)	95.6103
		RGB(3ch) + IR(3ch)	95.7709
		RGB(3ch) + IR(3ch) Mix	95.5567
RGB(3ch) + IR(3ch) Crop and Mix		93.0407	

[Table 2]는 연구에 사용된 컴퓨터 사양을 나타낸다. 위 사양과는 달리 사용된 분류 모델은 파라미터 수가 약 1600만 정도로 기존 Vision Transformer의 파라미터 수인 8600만에 비해 약 5.375배 작은 경량 모델이다.

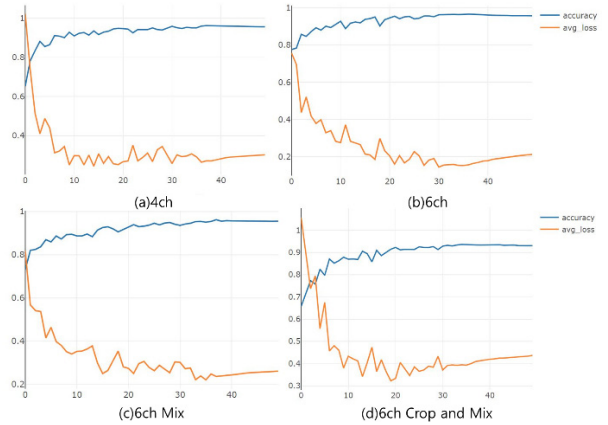
[Table 3]는 제안하는 방법으로 학습한 결과를 나타낸다. 학습에는 RGB와 IR 데이터가 필수로 사용되었으며, 학습 파라미터로는 50 Epoch와 20 Batch size, 0.001의 Learning late로 수행되었다. 비교를 위해 기본적인 방법인 RGB 데이터를 학습한 결과를 확인하였으며, 92.2912의 Accuracy를 보인다. 일반적으로 이미지 엔트로피 이론상 RGB와 IR의 결합 이후 정보량이 많아지는 것이 확인되는 것과 같이 전반적으로 Accuracy가 개선되는 것을 확인할 수 있다. Validation 데이터에도 IR 데이터가 포함되어 있으며, Accuracy 연산을 위하여 활용되기 때문에 어두운 환경에서 활용되는 IR 데이터 입력에도 반응하게 된다. 결과는 RGB 이미지 데이터만을 사용할 때보다 최대 3% 이상 개선되는 것이 확인된다.

[Fig. 6]는 RGB 데이터만을 학습시킨 Accuracy와 Loss 결과 그래프를 나타낸다. 논문에 사용된 자체 데이터셋의 경우 데이터가 작은 문제와 경량 학습 모델 구조로 인하여 초기 학습 시에 Loss 연산 값이 불안정한 모습을 보인다.

그에 반해 [Fig. 7]는 제안된 방법으로 수행된 4개의 결과를 보여준다. 이미지는 (a) ~ (d)까지 각각 IR 이미지를 1채널 활용한 4채널로 결합, IR을 3채널로 활용한 6채널 결합, 6채널을 섞은 후 결합, 6채널의 이미지 크기를 맞춘 후 결합하여 학습



[Fig. 6] Loss and accuracy graph during training on RGB data



[Fig. 7] Loss and accuracy graphs during the training of the proposed combined data learning

한 결과 그래프를 나타낸다. [Fig. 6]에 비해 학습에 사용되는 부족한 데이터를 정보량이 늘어남에 따라 Transformer 구조 활용에 따른 데이터의 양적 충당과 더불어 전반적인 학습에서 매우 안정적인 Loss 결과를 보여준다.

5. 결론

본 논문에서는 군 작전에 활용할 로봇의 인지 기술에 사용하기 위해 주야간의 이미지 데이터를 효과적으로 학습할 수 있도록 입력 데이터를 변형하는 방법을 연구하였다. 변형에는 기존 RGB 이미지와 동일 위치에서 취득되는 IR 이미지를 함께 활용하였으며, 군 작전에도 자주 볼 수 있는 데이터이다. 단순 이진 이미지 데이터만을 활용했을 때 자칫 부족할 수 있는 특징점을 색상 이미지와의 결합을 통해 보충하고, 야간 작전 시 IR 카메라를 통한 입력에도 반응할 수 있다. 사용되는 이미지는 연구에 활용된 적외선 감지 방식의 IR 뿐만 아니라 열 감지 방식의 데이터를 활용해도 무방할 것이다.

인공지능의 학습 시점에서 바라보면 실제 데이터를 통해 더욱 정확한 추가 데이터를 확보할 수 있게 되며, 한 개의 이미지는 기존보다 학습에 많은 의미 정보를 가지게 된다. 추후 해당 데이터 입력 방법을 적용하여 학습 모델 구조를 바꾸어 일부 성능을 높이거나, 데이터셋의 Labeling 작업과 Detection branch를 추가하여 객체 인식기능으로 변형 활용이 가능하다. 현재 로봇에 적용할 것을 목적으로 실내의 데이터를 위주로 활용하였으나, 드론과 야전에 적합한 로봇에도 활용될 수 있도록 야외 데이터의 확보와 테스트 수행이 필요하다.

References

[1] Ministry of National Defense, 2018 Defense white paper, 2019, [Online], https://www.mnd.go.kr/cop/pbictn/selectPublicationUser.do?siteId=mndEN&componentId=51&categoryId=0&publicationSeq=846&pageIndex=1&id=mndEN_031300000000, Accessed: May. 9, 2023.

[2] "Military innovation 4.0," *Ministry of National Defense*, [Online], https://www.mnd.go.kr/mbshome/mbs/mnd/subview.jsp?id=mnd_010302010000, Accessed: Mar. 3, 2023.

[3] K. S. Jang and Y. K. Cheung, "A study on the method acquiring NCOs according to the decrease in school age population -Activation of the department of navy noncommissioned officers," *Journal of Information and Security*, vol. 21, no. 1, pp. 159-168, Mar., 2021, DOI: 10.33778/kcsa.2021.21.1.159.

[4] C. Grandou, L. Wallace, H. H. K. Fullagar, R. Duffield, and S. Burley, "The Effects of Sleep Loss on Military Physical Performance," *Sports Med*, vol. 49, pp. 1159-1172, May., 2019, DOI: 10.1007/s40279-019-01123-8.

[5] C. H. Good, A. J. Brager, V. F. Capaldi, and V. Mysliwiec, "Sleep in the United States Military," *Neuropsychopharmacology*, vol. 45, pp. 176-191, Jun., 2020, DOI: 10.1038/s41386-019-0431-7.

[6] E. S. Kim, "Application of AI in Defense, U.S. Case Studies, and Key Considerations to Know," *KIDA Defense Issues & Analysis*, vol. 20, no. 18, pp. 1-9, May., 2020, [Online], <https://kida.re.kr/cmm/viewBoardImageFile.do?idx=28058>.

[7] Boannews, "CCTV Cameras for Military Border Security, Available Products," [Online], <https://www.boannews.com/media/view.asp?idx=101175>, Accessed: Oct. 5, 2021.

[8] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 million image database for scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1452-1464, Jul., 2017, DOI: 10.1109/TPAMI.2017.2723009.

[9] AXIS Communications, IR used for surveillance (White paper), Jun., 2018.

[10] "How do thermal camera work?," *FLIR Korea*, [Online], <https://www.flirkorea.com/discover/rd-science/how-do-thermal-cameras-work/>, Accessed: Jun. 16, 2020.

[11] F. Farahnakian, J. Poikonen, M. Laurinen, and J. Heikkonen, "Deep Convolutional Neural Network-based Fusion of RGB and IR Images in Marine Environment," *International Conference on Intelligent Transportation*, Auckland, New Zealand, 2019, DOI: 10.1109/ITSC.2019.8917332.

[12] S. Speth, A. Gonçalves, B. Rigault, S. Suzuki, M. Bouazizi, Y. Matsuo, and H. Predinger, "Deep learning with RGB and thermal images onboard a drone for monitoring operations," *Journal of Field Robotics*, vol. 39, no. 6, pp. 840-868, May., 2022, DOI: 10.1002/rob.22082.

[13] S. Dey, Hands-On Image Processing with Python, Sandipan Dey, 2018, [Online], <https://www.oreilly.com/library/view/hands-on-image-processing/9781789343731/>



최 동 규

2018 동의대학교 컴퓨터공학과(학사)
 2020 동의대학교 소프트웨어융합학과(석사)
 2023 동의대학교 컴퓨터공학과(박사)
 2023~현재 한국전자통신연구원 연구원

관심분야 인공지능, 이미지 처리, 객체분류/인식, 로봇제어, 상황/환경 인지



도 승 원

2018 포항공과대학교 전자전기공학과(학사)
 2020 한국과학기술원 문화기술대학원(석사)
 2020~현재 한국전자통신연구원 연구원

관심분야: 인간-컴퓨터 상호작용, 인공지능, 강화학습, 로봇제어



이 창 은

1996 한양대학교 전자공학과(학사)
 1998 한양대학교 전자공학과(석사)
 2017 충남대학교 정보통신공학과(박사)
 2001~현재 한국전자통신연구원 책임연구원

관심분야: 인공지능, 분산 및 협업로봇 제어, WPAN 기반 로봇 위치 추위