

한정된 군사 데이터를 활용한 이미지 분류 AI의 성능 향상 방안: Grad-CAM을 활용한 준지도학습 적용

정자훈¹, 김용기², 나성중³, 류준열^{4*}

¹육군대학, ²과학기술정책연구원, ³육군사관학교 기계·시스템공학과, ⁴국방기술진흥연구소

Enhancing Performance of Image Classification AI Using Limited Military Data: Application of Weakly Supervised Learning with Grad-CAM

Ja-Hoon Jeong¹, Yong-Gi Kim², Seong-Jung Na³, Jun-yeol Ryu^{4*}

¹Army College

²Science and Technology Policy Institute

³Department of Mechanical and Systems Engineering, Korea Military Academy

⁴Korea Research Institute for defense Technology Planning and Advancement

요약 무인 지상 차량, 무인 비행체 등과 같은 자율 무인체계에 탑재된 AI 모델은 센서를 통해 획득한 적 인원, 무기체계 등을 탐지하고 분류한다. 이때 무기체계를 정확하게 분류하는 것은 화력 및 장애물의 운용 등 작전 수행에 있어 중요한 사안이다. AI 모델의 성능 향상을 위해서는 적 인원, 무기체계에 대한 학습데이터가 필요하다. 그러나 평시 적 무기체계에 대한 이미지 데이터 등을 확보하기 어려울 뿐만 아니라 위장, 부착 무장의 변경 등의 다양한 요인으로 인해 전쟁 초기에는 평소 학습한 형태와 다른 상태의 적 무기체계를 분류해야 한다. 이 경우 초기에 확보된 부족한 적 무기체계 데이터를 학습하여 AI 모델의 분류성능을 향상해야 한다. 본 연구에서는 Grad-CAM을 활용하여 이미지 분류 모델이 학습한 데이터 영역을 분석하고, 분류를 위한 관심 구역에 맞춰 노이즈를 추가하는 준지도학습을 사용하여 적 무기체계에 대한 데이터가 부족한 상황에서도 AI 모델의 분류성능을 향상하는 방법을 제안하였다. 본 연구에서의 준지도학습을 적용했을 때 비교적 적은 수의 데이터를 학습시켜도 VGG-16과 MobilNetV2의 분류성능이 향상되는 것을 확인할 수 있었다. 향후 준지도학습을 적용하여 제한된 군사데이터 상황에서도 작전수행 역량을 향상시키는데 활용되기를 기대한다.

Abstract AI models embedded in autonomous unmanned systems classify enemy personnel and weapon systems acquired through sensors. Accurate classification of weapon systems is crucial for operational tasks. Training data on enemy weapon systems are required to enhance the performance of AI models. On the other hand, Acquiring image data during peacetime is challenging. During the initial stages of war, the AI model must classify enemy weapon systems in states different from what it was trained on because of factors such as camouflage and changes in attached armaments. In such cases, it is necessary to improve the classification performance of AI models by training them on the limited data of enemy weapon systems acquired at the early stages. In this study, Grad-CAM was utilized to analyze the data regions learned by image classification models. A Weakly Supervised Learning approach was proposed, which added noise to the regions of interest for classification, addressing situations with a shortage of data for enemy weapon systems. The classification performance of VGG-16 and MobileNetV2 improved even when trained on a relatively small amount of data. Weakly supervised learning can improve operational capabilities, even in limited military data.

Keywords : Autonomous Unmanned Systems, Grad-CAM, Weakly Supervised Learning, VGG-16, MobileNetV2, Limited Data

*Corresponding Author : Jun-yeol Ryu(Korea Research Institute for Defense Technology and Planning)
email: barampool@gmail.com

Received August 9, 2023

Revised August 31, 2023

Accepted September 1, 2023

Published September 30, 2023

1. 서론

국방혁신 4.0과 미래 육군의 발전 방향을 제시한 ARMY TIGER 4.0에 따르면, 향후 전장에서 무인 지상 차량, 무인 비행체와 같은 무인 전투체계가 다수 활용될 것이다[1-3]. 초기에는 원격 통제에 기반하여 개발되고 향후에는 반자율, 자율 형태로 개발된다[4]. 최종 의사결정은 무인체계를 통제하는 운용관이 하지만 전투 상황에서 전투 피로의 가중, 임무 부하로 인한 인지능력의 한계 등으로 인해 실시간 수집되는 다량의 정보를 운용관이 모두 판단하기 제한된다. 따라서 운용관에게 정보를 제공하기 전 AI 모델 등이 분석한 정보를 제공하는 것에 대한 중요성이 높아지고 있다. 무인체계는 EO/IR 등의 센서를 사용하여 감시 및 정찰 임무를 수행한다. 이때 중요한 것은 식별된 표적의 종류를 신속하게 분류하여 무인체계를 통제하는 운용관에게 정보를 제공하여 적절한 대응이 가능하도록 하는 것이다. 적 무기체계에 대한 정확한 분류를 위해 사전에 확보한 적 무기에 대한 이미지, 영상 등의 데이터를 활용하여 무인체계에 탑재된 AI 모델을 훈련한다. 그러나 공개된 자료가 부족하여 실전에 투입되기 전 AI 모델을 충분히 학습시키기 어려울 뿐만 아니라 위장, 외부 부착 센서 및 화기의 다양성 등으로 인해 적 무기체계의 분류에 변수가 발생한다. 그러므로 전쟁 초기는 제한적으로 확보된 적 무기체계 정보에 기반하여 AI 모델을 훈련하게 된다. 초기 수집된 데이터의 양은 제한적이며 이는 무인체계가 사용하는 AI 모델의 적 무기체계 분류 신뢰성을 저하한다. AI 모델이 적 무기체계 분류를 정확하게 하지 못할 경우 작전 수행 능력에 문제를 야기할 수 있다. 예를 들어 주포, 조준경 등으로 인해 전차와 외관이 유사한 IFV(Infantry Fighting Vehicle)를 전차로 분류하면 방책 선정 등에 있어 기동로 선정, 장애물 배치, 화력계획 등의 작전 구상에 있어 잘못된 의사결정을 초래할 수 있다. 따라서 전쟁 초기에 적 무기체계에 대한 데이터가 부족한 상황에서 AI 모델의 분류 정확성을 높이는 것은 중요한 문제이다.

본 연구에서는 실전 상황의 데이터를 최대한 활용하기 위해 러시아-우크라이나전에서 수집한 기갑 및 기계화 무기체계 이미지 데이터를 주로 활용하였다. 실상황과 유사한 전례를 바탕으로 데이터가 부족한 상태에서 모델의 성능을 향상하기 위해 Grad-CAM을 사용하여 연구를 진행하였다. CNN 기반의 AI가 이미지를 클래스별로 분류할 때 분류의 기준이 되는 이미지 영역(R.O.I. / Regions Of Interest)을 Grad-CAM을 통해 확인하고, 오분류된

이미지 중 잘못된 ROI가 선정된 학습데이터를 수정함으로써 재학습 간 무기체계 분류에 대한 정확도를 향상할 수 있었다. 이는 향후 적 무기체계에 대한 데이터가 부족한 상황에서 설명할 수 있는 AI를 통한 준지도학습(Weakly Supervised Learning)으로 무인체계를 효과적으로 운용하는 데 유용한 역할을 할 것으로 기대한다.

본 논문은 총 4장으로 구성되어 있다. 2장에서는 관련 연구를 살펴보고 연구 방법 및 데이터 구성에 대해 고찰한다. 3장에서는 실험설계 및 결과를 분석한다. 4장에서는 결론 및 향후 연구 방향에 대해 제시한다.

2. 연구 방법 및 데이터 구성

2.1 관련 연구

인공지능이 활발하게 사용되고 있는 이미지 분류 분야에서는 딥러닝(Deep Learning)의 분류 성능이 기계학습(Machine Learning)에 비해 우수한 성과를 보여주어 주요 AI 기법으로 적극 활용되고 있다[5]. 분류 성능을 높이기 위해서는 충분한 양의 데이터가 필요하다. 이미지 분류 분야에서 널리 활용되는 CNN(Convolutional Neural Network)모델도 부족하게 수집된 학습 데이터는 해당 AI 모델의 분류성능을 저하해 신뢰성 있는 결과를 도출하는데 제한이 있다[6].

앞서 언급한 바와 같이 불충분한 학습용 이미지 데이터의 한계를 극복하고 AI의 예측 성능을 향상하기 위한 여러 가지 방안들이 제시되고 있다[7]. 그 중 대표적인 것은 ①데이터 증강(Data Augmentation)과 ②전이 학습(Transfer Learning)이다. ①데이터 증강은 이미지 데이터를 변형시켜 추가적인 학습데이터를 생성하는 기법으로, 회전, 이동, 반전, 밝기 조절 등의 변형을 적용하여 한정된 데이터의 다양성을 높인다는 장점이 있지만, 원본 데이터에서 변형을 가하는 것이기에 단순하게 적용하면 딥러닝 모델의 과적합(Overfitting)을 야기한다는 한계를 가지고 있다[8]. ②전이 학습은 미리 학습된 모델을 이용하여 새로운 작업에 활용하는 기법으로, 이미지넷(Image-Net)과 같은 대규모 데이터에서 학습된 모델을 다른 작업에 재사용하는 방법을 말한다. 이때, 해당 모델에 저장된 초기 가중치를 그대로 사용하거나 미세조정(Fine Tuning)을 통해 분석하고자 하는 데이터를 더 잘 예측할 수 있다. 하지만, 기존 학습된 데이터들에 존재하지 않는 도메인이 불일치하는 경우, 또는 기존 모델에서 잘 분류해 내지 못한 도메인에 대해 분석하는 경우

그 예측력이 충분치 못할 수 있다는 한계가 있다.

이처럼 데이터양의 부족 문제를 해결하고자 하는 모델들의 한계를 극복하고자 본 논문에서 제시하는 방법은 Grad-CAM을 활용한 준지도학습 방법이다. Grad-CAM은 Gradient-weighted Class Activation Mapping의 약자로, 설명할 수 있는 AI(eXplainable AI / XAI) 중 하나의 기법이다[9]. 이미지를 분류하는 CNN 기반의 AI 모델에서 특정 클래스를 분류할 때 중요하게 고려한 영역을 표시해 주는 것으로, 사람이 이미지를 분류해 낼 때 거치는 과정과 유사하다.

Grad-CAM은 CAM(Class Activation Map)의 단점을 보완하여 나온 개념이다. 먼저 CAM이란, CNN 기반 모델의 마지막 합성곱 신경망 층 FC(Fully Connected Layer)을 전달하는 과정에서 GAP(Global Average Pooling)을 사용하는 경우 정보의 손실이 줄어드는 점에 착안해서 나온 개념으로, 최종 출력 클래스와 GAP 층 사이의 가중치와 마지막 합성곱 층의 출력인 특징맵과의 곱으로 계산되며 이를 수식으로 표현하면 아래와 같다[10].

$$M_c(x, y) = \sum_k w_k^c f_k(x, y)$$

이를 통해 계산된 값들을 Heat Map화 하면 CNN 모델이 어떤 픽셀들에 가중치를 두고 이미지를 분류했는지 알 수 있다. 하지만, CAM의 단점은 출력층과 마지막 합성곱 층 사이의 층이 GAP를 사용할 때만 연산이 가능하다는 한계를 가지고 있다. 이를 보완하기 위해 마지막 전 연결층이 존재하더라도 출력에 대한 특정 층의 Gradient를 계산하면 CAM을 얻을 수 있었고, 이를 Grad-CAM이라고 한다[11].

아래 Fig. 1은 Grad-CAM의 예시를 보여준다. Fig. 1의 왼쪽 (a)는 원본 이미지이고 (b)는 CNN 모델을 활용하여 AI가 해당 이미지의 클래스를 예측할 때 어떤 부분에 중점을 두었는지를 Heat Map을 활용해 전시한 것이다. (b)에서 밝게 표시된 전차의 배연기, 포탑 전면부의 감시 기재 등이 클래스 분류에 높은 영향을 주었음을 알 수 있다.



Fig. 1. An example of Grad-CAM
 (a) An Original Image
 (b) The Result of Grad-CAM

위와 같은 Grad-CAM의 결과물을 활용하여, 이미지 분류 AI의 성능을 향상하는 여러 방법이 활발히 연구되고 있으며[12-14], 본 논문에서는 모델의 오분류 원인을 확인하고 AI가 중점을 잘못 둔 경우, 다시 말해 이미지 내에서 분류해야 할 객체가 아닌 주변 환경에 대해 주목하는 경우들을 파악하고, 추후 AI의 학습 간에는 해당 부분에 중점을 두지 못하도록 Noise를 추가하여 분류해야 하는 객체에 관심을 가지도록 하는 준지도학습을 제안한다.

2.2 분석 절차

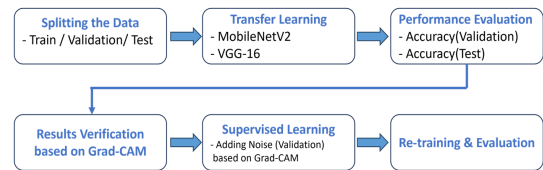


Fig. 2. Analysis Process

본 논문에서의 분석 절차는 위 Fig. 2와 같으며 단계별 세부 사항은 아래에 기술하였다.

- ① 데이터 분류 : 각 6개의 클래스의 전체 이미지 데이터를, 학습용 데이터(60%)와 검증용 데이터 (20%), 그리고 테스트 데이터(20%)로 각각 구분한다.
- ② 전이 학습 : MobileNETV2와 VGG-16을 기본 모델로 하여 학습 데이터에 대해 전이학습을 실시하여 그 성능을 평가한다.

MobileNetV2은 2018년도 Mark Sadler 등 5명의 저자에 의해 제안된 컴퓨터 비전 학습 모델이다. 이는 모바일과 같은 컴퓨팅이 환경이 떨어지는 환경에서도 원활하게 동작할 수 있는 연산량 및 파라미터 수가 적지만 높은 분류 정확도를 유지하는 모델이다. 이를 위해 Depthwise Separable Convolution과 Inverted Residual이라는 개념을 통해 연산량 및 파라미터 수를 줄이면서도 모델의 분류성능을 유지하였다[15]. VGG-16은 1,000개의 클래스로 구성된 약 1,400만 개의 이미지 데이터를 분류하는 대외인 ILSVRC(ImageNet Large Scale Visual Recognition Challenge/이미지인식 경진대회)에서 92.7%의 Top-5의 테스트 정확도를 달성하며 2014년에 준우승한 모델로 Simoyan과 Zisser에 의해 개발되었다. 이는 16층의 깊은 신경망에도 불구하고 3x3 합성곱 필터를 사용함으로써 상대적으로 적은 학습 파라미터 수를 도출함과 동시에 보다 높은 비선형성을 제공하여 시간 및 정확도 측면에서 효율적인 모델로 평가받고 있다[16].

학습 시간 단축을 위해 각 모델이 ImageNet 데이터를 통해 학습된 가중치 값은 그대로 활용하며, 마지막 층에 전연결층을 추가하여 6개의 클래스로 구성된 연구 데이터들이 분석될 수 있도록 한다. 이때, 총학습 횟수는 빠른 학습 속도와 성능이 준수한 MobileNetV2모델의 학습 결과를 기준으로 30회로 선정하였으며, 이는 해당 모델이 최초 학습 간 Validation Loss가 증가하는 시점(과적합 발생 시점)인 20회에 여유 횟수인 10회를 추가한 것이다.

- ③ 성능 확인 : ②를 통해 각 모델이 도출하는 테스트 데이터의 정확도(Test Set Accuracy)와 검증용 데이터의 정확도(Validation Set Accuracy)를 확인한다. 또한, 각 데이터(검증용, 테스트용)에 대해 Grad-CAM을 활용한 결과물을 도출한다.
- ④ Grad-CAM을 활용한 결과 확인 : 검증용 데이터의 분류 결과를 Grad-CAM을 활용한 Heat Map으로 도출한 뒤 각 이미지 분류 모델들이 분류 간 중점을 둔 부분들을 확인한다. 이때, 분류하고자 하는 기갑 및 기계화 전력이 포함된 이미지 영역이 아닌 기타 영역(주변의 하늘, 나무, 건물 등등)에 중점을 둔 데이터들을 분류해 낸다.
- ⑤ 준지도학습(Noise 추가) : ④에서 분류한 이미지 데이터들(검증용 데이터 중 오분류 되었으며, 관심 영역이 잘못 설정된 데이터들)에 대해 Noise를 추가한다. Noise를 추가하는 영역은 Grad-CAM의 결과 중 기갑 및 기계화 전력이 포함된 이미지 영역이 아닌 기타 영역이며, 해당 영역에 정규화된 Noise를 추가함으로써 이미지 모델이 학습 간 해당 부분에 중점을 두지 못하도록 한다. 이렇게 Noise가 추가된 검증용 데이터들은 원본 검증용 데이터들을 대체하여 분석에 활용한다.
- ⑥ 재학습 및 평가 : ⑤에서 재설정된 검증용 데이터를 바탕으로 재학습을 수행하고 그 분류 성능을 테스트 데이터의 분류 정확도를 기반으로 평가한다.

2.3 데이터 구성

본 연구에서는 공개적으로 접근할 수 있는 우크라이나 전 데이터를 주로 사용하여 분석하였다. 데이터의 수집 방법으로는 러시아 및 우크라이나의 국방부 홈페이지와 CNN, BBC 등 유력 언론에 보도된 자료, 다국적군이 수집하여 공개한 자료, 북한이 사용하는 러시아 계열 장비 등 대중에게 공개된 자료를 활용하였다. 이와 같은 방법으로 수집한 자료는 총 6개의 클래스(Class)이며, 각 클

래스는 ①BMP ②BTR ③T54 ④T55 ⑤T62 ⑥T72이다. 장갑차는 궤도형인 BMP와 차륜형인 BTR로 구성되어 있으며, T계열 전차(T-54, T-55, T-62, T-72)는 기종별 무장, 방호력에 차이가 있고 기종별 대응 방법이 상이하므로 세부적인 클래스 구분이 필요하므로 시리즈별로 구성하였다. 기종별 이미지 데이터는 아래 Table 1과 같다. Table 1에 명시된 것처럼 각 기갑 및 기계화 전력의 클래스별 데이터의 숫자는 충분하지 않다. 이처럼 데이터가 부족할 경우 CNN과 같은 딥러닝 모델의 성능은 저하된다. 충분한 데이터가 없으면 AI 모델은 분석하고자 하는 이미지의 공통적인 특정 패턴을 학습하지 못하거나 지나치게 데이터의 특정 부분에 의존하여 학습하게 되어 예측 능력이 저하되는 문제가 발생한다. 이는 앞서 언급한 바와 같이 적 무기체계를 오분류하여 작전 수행역량을 저하하는 결과를 야기한다. 이러한 문제점을 해결하기 위해 본 연구에서는 Grad-CAM을 활용한 준지도학습을 적용하였다.

Table 1. Classes of the Data Set

Image		
Class	BMP	BTR
# of Data	71	55
Image		
Class	T-54	T-55
# of Data	27	38
Image		
Class	T-62	T-72
# of Data	55	80

3. 실험 설계 및 결과

3.1 전이학습 모델 설정

분류에 사용된 MobileNetV2와 VGG-16은 각각 다음 페이지 Fig. 3과 Fig. 4의 구조를 가진 모델들이다. 이때, 해당 모델들은 1,000개의 클래스를 가진 ImageNet 데이터들을 바탕으로 학습이 되었기에 마지막 결과를 도출하는 전연결층은 1,000개의 클래스가 구분되도록 설

정되어있다. 본 연구에서 활용한 데이터는 총 6개의 클래스를 가졌기 때문에, 각 모델이 기존에 가지고 있었던 전 연결 층을 제거하고 앞서 언급한 6개의 클래스가 구분되도록 새로운 연결 층을 생성한다. 이때, 각 모델이 ImageNet을 통해 사전 학습된 가중치들을 활용한다.

두 모델 모두 연구수행 과정에서 확인한 Loss 값을 고려하여 30번의 학습을 하였으며, 이때 검증용 데이터의 Loss가 5번 연속 증가하게 된다면 학습을 중단시키는 Early Stopping을 설정하였다.

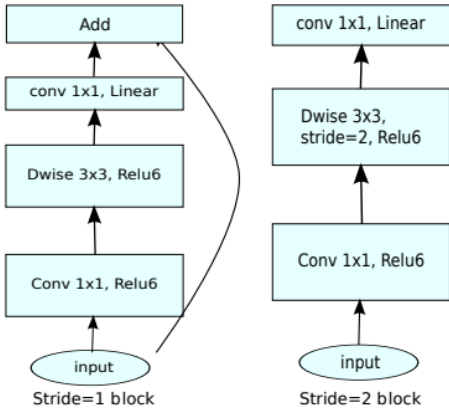


Fig. 3. The Architecture of MobileNetV2

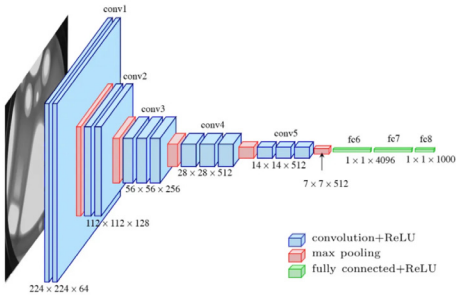


Fig. 4. The Architecture of VGG-16

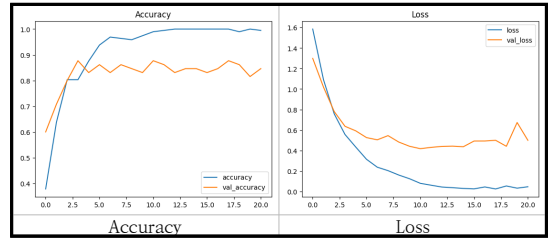
3.2 최초 분류성능 확인

최초 분류성능은 ①학습용 및 검증용 데이터를 활용해 최적의 모델을 선정하고, ②테스트용 데이터를 활용해 선정된 모델의 분류성능을 확인하는 과정이다.

3.2.1 MobileNetV2

마지막 층(Layer)에 6개의 클래스 분류를 위한 전연결층을 추가한 MobileNetV2의 학습 및 검증 결과는 오른쪽 Table 2와 같다.

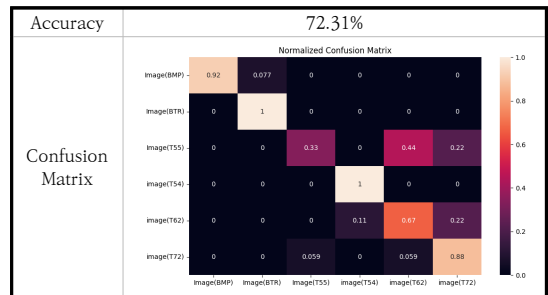
Table 2. Results derived from Train & Validation Data



위 두 그래프에서 하늘색 선은 학습용 데이터를 주황색 선은 검증용 데이터의 분류 결과를 나타낸다. 모델의 학습은 최초 30번을 기준으로 설정하였으나, 15번째 학습 이후부터 검증용 데이터의 Loss가 이후 5번의 학습 간 증가하게 되어 학습을 중단하고 20회차의 모델을 최적의 모델을 선정하였다. 해당 모델(20회차)로 검증 세트를 분류했을 경우 83%(=0.83)라는 분류 정확도가 도출되었다.

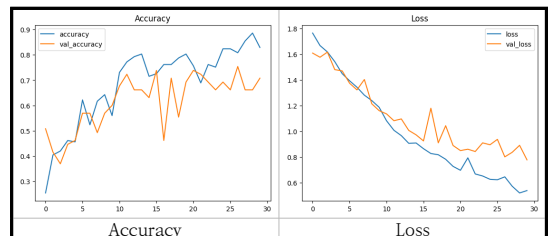
이 모델로 테스트 데이터를 통해 분류한 결과는 아래 Table 3과 같다. 전연결층을 추가한 MobileNetV2의 분류 정확도는 72.31%이다. BMP, BTR, T-54, T-72는 88% 이상의 정확도로 분류되었으나, T-62는 67%의 분류 정확도를 보였으며 T-55의 경우에는 33%라는 매우 낮은 분류 정확도를 보였다.

Table 3. Results derived from Test Data



3.2.2 VGG-16

Table 4. Results derived from Train & Validation Data

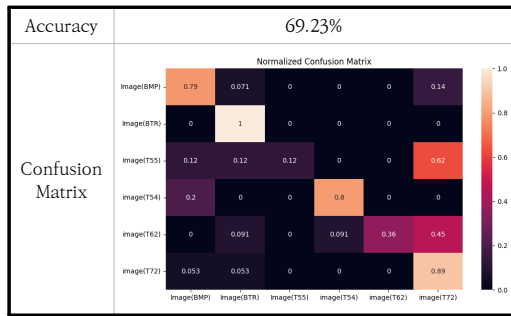


마지막 층(Layer)에 6개의 클래스 분류를 위한 전연 결층을 추가한 VGG-16의 학습 및 검증 결과는 앞 페이지 Table 4와 같다.

모델의 학습은 최초 30번을 기준으로 설정하였으며, 28회차의 학습 간 발생한 검증용 데이터의 Loss가 가장 낮기 때문에 해당 회차(28회)의 모델을 최적의 모델로 선정하였다. 해당 모델(20회차)로 검증 세트를 분류했을 경우 75.38%(=0.7538)라는 분류 정확도가 도출되었다.

앞서 도출된 최적의 모델로 테스트 데이터를 분류한 결과, 분류 정확도는 아래 Table 5에 명시된 바와 같이 69.23%이다. BMP, BTR, T-54, T-72는 약 80% 이상의 정확도로 분류되었으나 T-62는 36%, T-55의 경우에는 12%라는 매우 낮은 분류 정확도를 보였다.

Table 5. Results derived from Test Data

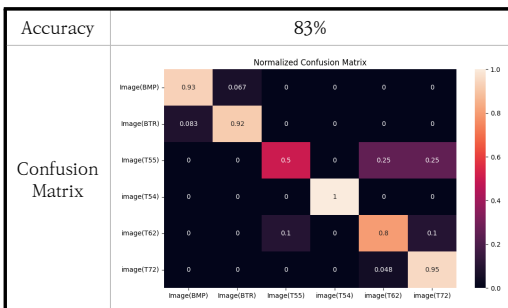


3.3 Grad-CAM을 활용한 결과 확인

이 과정은 설명할 수 있는 AI(Explainable AI) 기법의 하나인 Grad-CAM을 활용하여 모델 학습간 사용된 검증용 데이터의 분류 결과를 해석하는 과정이다.

3.3.1 MobileNetV2

Table 6. Results derived from Validation Data



MobileNetV2를 활용하여 도출된 최적의 모델의 경우, 검증용 데이터 총 65개 중 11개의 데이터가 오분류 되어 위 Table 6과 같이 83%의 분류 정확도를 보였다.

MobileNetV2가 검증데이터를 분류 간 오분류된 데이터는 아래 Table 7에 Heat Map으로 표현된 것과 같이 기갑 및 기계화 전력이 표현된 영역이 아닌 주변 영역에 중점을 두고 분류했음을 알 수 있다.

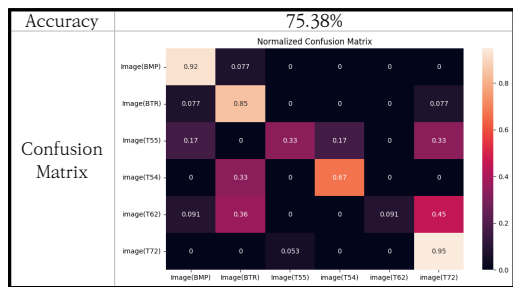
Table 7. Heatmap derived from Grad-CAM (Validation Data)



3.3.2 VGG-16

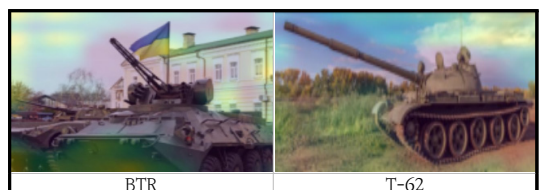
VGG-16을 활용하여 도출된 최적의 모델의 경우, 검증용 데이터 총 65개 중 16개의 데이터가 오분류 되어 아래 Table 8과 같이 75.38%의 분류 정확도를 보였다.

Table 8. Results derived from Validation Data



VGG-16이 검증 데이터를 분류 간 오분류된 데이터는 아래 Table 9와 같이 기갑 및 기계화 전력의 주포, 감시 기재 등의 특징적인 영역이 아닌 주변 영역에 중점을 두고 분류했음을 알 수 있다.

Table 9. Heatmap derived from Grad-CAM (Validation Data)



3.4 준지도학습(검증용 데이터에 Noise 추가)

3.4.1 MobileNetV2

앞서 살펴본 Grad-CAM을 통해 도출된 Heat Map을 확인한 결과, MobileNetV2가 오분류한 검증용 이미지의 가중치가 높은 관심 영역은 전차 또는 장갑차가 포함된 영역이 아닌 주변 영역이었다.

아래 Table 10의 예시와 같이 해당 영역에 정규화된 Noise를 추가함으로써 이후 학습 간에 해당 영역이 분류 성능에 미치는 영향력(가중치)을 낮추도록 처리하였다. 즉, 해당 영역을 학습에서 제외해 분류 모델이 중점이 되는 이미지의 학습에 더 중점을 주도록 하는 준지도학습을 구현하였다.

Table 10. Heatmap derived from Grad-CAM (Validation Data)



3.4.2 VGG-16

앞서 살펴본 Grad-CAM을 통해 도출된 Heat Map을 확인한 결과, VGG-16이 오분류한 검증용 이미지의 가중치가 높은 관심 영역은 MobileNetV2와 같이 전차나 장갑차가 포함된 영역이 아닌 배경 영역이었다. 아래 Table 11과 같이 해당 영역에 정규화된 Noise를 추가하여 무기체계의 특징적 부분에 집중하여 분류할 수 있도록 준지도학습을 구현하였다.

Table 11. Heatmap derived from Grad-CAM (Validation Data)



3.5 재학습 및 평가

총 65개의 검증용 데이터 중 오분류된 데이터를 도출

하고, 그중 기갑 및 기계화 전력이 포함된 영역이 아닌 주변 영역을 중점으로 분류한 데이터들에 정규화된 Noise를 추가하였다. 이후 해당 이미지들은 기존 이미지들을 대체하여 검증용 데이터에 포함하였고, 수정된 검증용 데이터를 바탕으로 재학습 및 평가를 하였다.

3.5.1 MobileNetV2

MobileNetV2를 기본 모델로 선정한 경우, 총 65개의 검증용 데이터 중 7개가 오분류 되었고, 그중 3개의 데이터가 주변 영역에 높은 가중치를 두고 있었다. 해당 이미지들에 Noise를 추가하고 재학습한 결과는 아래 Table 12와 같다.

아래 결과는 수정된 이미지들 검증용 데이터에 추가하여 분류된 결과이다.

Table 12. Results of Weakly Supervised Learning (Validation Data)



Validation Set의 Grad CAM을 확인해 보니, 기존에는 기갑 및 기계화 전력이 포함되지 않았던 영역에 ROI가 설정되었으나, Noise 추가를 통한 준지도학습을 실시하여 이전보다 기갑 및 기계화 전력이 포함된 영역에 중점을 두고 학습하였음을 확인할 수 있었다.

아래 Table 13과 다음 페이지 Table 14는 수정된 모델을 통해 테스트용 데이터를 분류한 결과이다.

Table 13. Results derived from Test Data(Before Applying Weakly Supervised Learning)

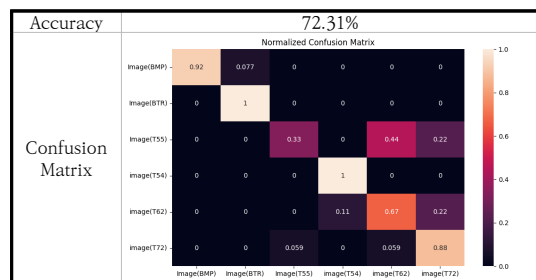
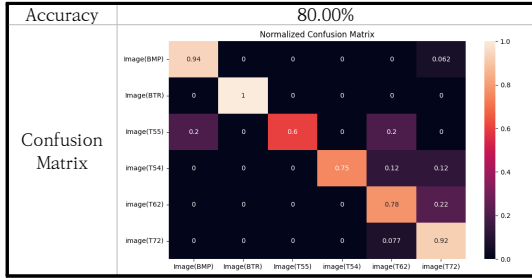


Table 14. Results derived from Test Data(After Applying Supervised Learning)



기존에는 테스트 데이터의 분류 결과가 72.31%였으나, Grad-CAM 결과를 통해 검증용 데이터의 오분류된 이미지에 Noise를 추가하고 재학습 시킨 경우 테스트 데이터의 분류 결과가 80%로 상승하였다.

3.5.2 VGG-16

VGG-16을 기본 모델로 선정한 경우, 총 65개의 검증용 데이터 중 16개가 오분류 되었고, 그중 5개의 데이터가 주변 영역에 높은 가중치를 두고 있었다. 해당 이미지에 Noise를 추가하고 재학습한 결과는 아래 Table 15와 같다.

Table 15. Results of Weakly Supervised Learning (Validation Data)



아래 Table 16과 오른쪽 Table 17은 수정된 모델을 통해 테스트용 데이터를 분류한 결과이다.

Table 16. Results derived from Test Data(Before Applying Weakly Supervised Learning)

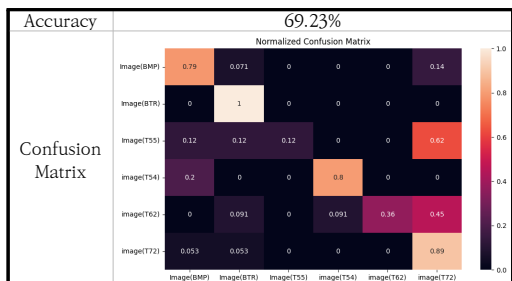
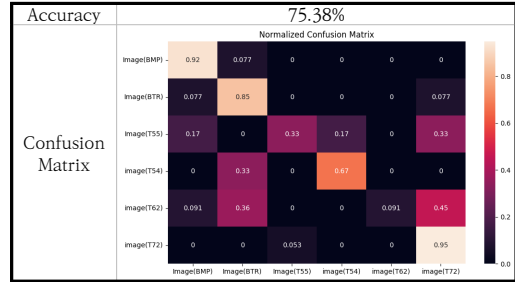


Table 17. Results derived from Validation Data (After Applying Weakly Supervised Learning)



기존에는 테스트 데이터의 분류 결과가 69.23%였으나, Grad-CAM 결과를 통해 검증용 데이터의 오분류된 이미지에 Noise를 추가하고 재학습 시킨 경우 테스트 데이터의 분류 결과가 75.38%로 상승하였다. 두 모델 (MobileNetV2와 VGG-16)을 기반으로 Grad-CAM을 활용한 준지도학습을 적용한 결과는 아래 Table 18과 같이 정리할 수 있으며, 두 모델 모두 준지도학습을 통해 분류 성능이 향상되었음을 확인할 수 있다.

Table 18. Comparison of Test Data Classification Performance Results

CNN Model	MobileNetV2	VGG-16
Test Accuracy (Before Applying W.S.L.)	72.31%	69.23%
Test Accuracy (After Applying W.S.L.)	80%	75.38%

*W.S.L. : Weakly Supervised Learning

Grad-CAM을 사용하여 준지도학습을 적용하면, 적은 수의 무기체계 데이터를 사용하여도 모델의 분류성능을 향상할 수 있다. 적 무기체계에 대한 데이터는 평시에도 확보가 어렵기 때문에 전쟁 초기 확보된 데이터를 효과적으로 학습할 방법이 필요하다. 본 연구에서 제안하는 방법을 적용하면 지상 무기체계뿐만 아니라 무인 비행체, 무인 수상정 등에서도 활용이 가능하여 무인체계의 임무 수행 능력을 향상할 수 있을 것으로 기대한다.

4. 결론 및 향후 연구 방향

이번 연구에서 제안한 방법은 AI가 도출한 결과물에 대해 지도자의 개입을 통한 전처리가 필요하다는 단점이

존재하나, 적은 수의 데이터의 한계를 극복하고 AI의 분류 정확도를 향상할 수 있다는 점에서 효과적이다. 특히 전쟁 초기와 같이 가용한 적 무기체계에 대한 데이터가 어려운 상황에서 유용하게 사용할 수 있는 방법이다.

전쟁 초기와 같이 획득되는 데이터가 부족한 경우 본 연구에서 사용한 설명 가능한 AI(Explainable AI) 중 하나인 Grad-CAM을 활용해 준지도학습을 한다면 크게 두 가지 이점이 있음을 확인하였다. 첫째는 학습 결과를 운용관이 시각적으로 이해하고 평가한 뒤, AI 모델 성능 향상을 위해 지도할 수 있다는 것이다. 둘째는 기존 데이터 증강과 같은 방법(Data Augmentation, Fine Tuning 등)이 가지고 있는 학습데이터에 대한 의존성 문제를 일부 해결할 수 있음을 확인하였다.

본 연구에서 제안한 방법은 두 가지 제한점이 있다. 첫째는 데이터 부족으로 인해 발생하는 과적합 문제가 있으며, 둘째는 학습/검증/테스트 데이터의 구성에 따라 도출되는 결과의 불안정성 문제를 근본적으로 해결하기 제한된다는 것이다.

향후에는 다양한 형태의 데이터를 확보하여 과적합 문제를 해결하는 방법에 대한 연구와 세트 구성에 영향이 적은, 성능적 강건성이 높은 모델의 연구가 필요하다. 또한 다양한 준지도학습을 적용했을 때, 모델의 성능에 미치는 영향 대한 추가 연구가 필요하다.

References

- [1] Ministry of Defense(MnD), Republic of Korea "Defense Innovation 4.0.", p.11-24, MnD, Republic of Korea., 2023.
- [2] Republic of Korea Army, Army Tiger 4.0, Available From: <https://www.army.mil.kr/army/12/subview.do> (accessed 2023 Aug 1)
- [3] J. Kim, S. Park, J. Cha, Y. Kim, "Future tactical communication system development plan through Army TIGER information distribution capability analysis", *KOCOSA*, Vol.21, No.4, pp.23-30, 2021. DOI: <https://doi.org/10.33778/kcsa.2021.21.4.023>
- [4] H. Jeong, T. Jeong, J. Kim, Y. Baek, W. Jeon, "Evaluation of technology level of MUM-T", p.25, Korea Research Institute of Technology and Planning, 2022, p.2-11
- [5] B. Kim, "Deep learning-based image classification technology trends", *The Journal of The Korean Institute of Communication Sciences*, Vol.35, No.12, pp.8-14, 2018.
- [6] W. Jung, D. Jeong, Y. Kim, C. Kim, H. Lee, H. Yu, J. Ryu, H. Oh, "Deep Generative Models to Overcome an Insufficient Data Problem in Structural Health Diagnosis", *Transaction of the Korean Society of Mechanical Engineers*, Vol.43, No.3, pp.169-176, 2019 DOI: <https://doi.org/10.3795/KSME-A.2019.43.3.169>
- [7] Moein, M. M., Saradar, A., Rahmati, K., Mousavinejad, S. H. G., Bristow, J., Aramali, V., & Karakouzian, M., "Predictive models for concrete properties using machine learning and deep learning approaches: A review", *Journal of Building Engineering*, Vol.63, Part A, 105444, 2023 DOI: <https://doi.org/10.1016/j.jobbe.2022.105444>
- [8] C. Park, B. Kim, I. Kim, J. Gwak, "Style-Generative Adversarial Networks for Data Augmentation of Human Images at Homecare Environments", *Korea Information Processing Society*, Vol.29, No.2, pp.565-567, 2022
- [9] R. Selvaraju, A. Das, R. Vedantam, M. Cogswell, D. Parikh, D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization", *International Convergence of Computer Vision*, 2019 DOI: <https://doi.org/10.48550/arXiv.1610.02391>
- [10] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, "Learning Deep Features for Discriminative Localization", *Institute of Electrical and Electronics Engineers*, pp.2021-2929, 2016 DOI: <https://doi.org/10.48550/arXiv.1512.04150>
- [11] R. Selvaraju, A. Das, R. Vedantam, M. Cogswell, D. Parikh, D. Batra, "Grad-CAM: Visual Explanation from Deep Networks via Gradient-Based Localization", *International Conference on Computer Vision*, pp.618-626, 2017 DOI: <https://doi.org/10.1007/s11263-019-01228-7>
- [12] S. Kang, "CAMICAP : A Study on the Performance Improvement Algorithm of RICAP-Based Data Augmentation Techniques Using Grad-CAM", *Korean Institute of Intelligent Systems*, Vol.32, No.4, pp.299-306, 2022 DOI: <https://doi.org/10.5391/JKIIS.2022.32.4.299>
- [13] S. Kim, J. Lee, N. Gwak, S. Ryu, J. Ahn, "Grad-CAM based deep learning network for location detection of the main object", *Korea Institute of information and Communication Engineering*, Vol.24, No.2, pp.204-211, 2020 DOI: <https://doi.org/10.6109/jkiice.2020.24.2.204>
- [14] S. Sim, C. Heo, S. Kang, "A Transfer Information Optimization Method through the Generation of Noise Based on Grad CAM", *Korean Institute of Intelligent Systems*, Vol.31, No.1, pp.82-87, 2021 DOI: <https://doi.org/10.5391/JKIIS.2021.31.1.082>
- [15] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", *The IEEE Conference on Computer Vision and Pattern Recognition*, Vol.4, pp.4510-4520, 2019. DOI: <https://doi.org/10.48550/arXiv.1801.04381>
- [16] J. Ryu, T. Kim, J. Jeong, "An AI Model for Classifying The State of Armored and Mechanized Weapon Systems Based on VGG16", *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 23, No.12, pp.741-748, 2022 DOI: <https://doi.org/10.5762/KAIS.2022.23.12.741>

정 자 훈(Ja-Hoon Jeong)

[정회원]



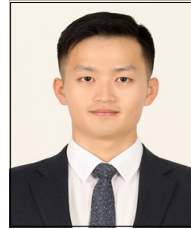
- 2013년 2월 : 육군사관학교 토목 환경공학과 (공학사, 군사학사)
- 2021년 8월 : 노스캐롤라이나 주립대학교 산업시스템공학 (공학석사)
- 2021년 8월 ~ 2022년 11월 : 육군사관학교 기계·시스템공학과 조교수
- 2022년 11월 ~ 현재 : 육군대학

<관심분야>

무기체계, 국방M&S, 데이터 분석, AI

류 준 열(Jun-Yeol Ryu)

[정회원]



- 2014년 2월 : 육군사관학교 무기 시스템공학과 (공학사, 군사학사)
- 2020년 2월 : 서울대학교 산업공학과 (공학석사)
- 2020년 1월 ~ 2023년 2월 : 육군사관학교 기계·시스템공학과 조교수
- 2023년 3월 ~ 현재 : 국방기술진흥연구소 연구원

<관심분야>

무기체계 설계, Robotic Combat Vehicle, M&S, AI

김 용 기(Yong-Gi Kim)

[정회원]



- 2007년 8월 : 포항공과대학교 산업경영공학과 (공학사)
- 2010년 5월 : KDI국제정책대학원 (경영학 석사, 정책학 석사)
- 2022년 2월 : 서울대학교 환경대학원 (환경 계획학 박사)

- 2022년 7월 ~ 현재 : 과학기술정책연구원, 글로벌 혁신전략연구 본부, 부연구위원

<관심분야>

인공지능 반도체, AI, 글로벌 산업 다이내믹스, 데이터 사이언스

나 성 중(Seong-Jung Na)

[정회원]



- 2015년 3월 : 육군사관학교 무기 시스템공학과 (공학사, 군사학사)
- 2023년 3월 : 국방대학교 무기체계 석사 (공학석사)
- 2023년 3월 : 미국 해군대학원 시스템공학 석사 (공학석사)
- 2023년 3월 ~ 현재 : 육군사관학교 기계·시스템공학과 강사

<관심분야>

무기체계, 국방M&S, 최적화