

지식증류 방안을 활용한 무인 군사 이미지 분류 AI 모델의 데이터 부족 및 경량화 모델 한계 극복

정자훈¹, 송윤호¹, 강인욱², 류준열^{3*}

¹육군사관학교 기계시스템공학과, ²국방기술진흥연구소, ³서울대학교 산업공학과

Addressing Data Scarcity and Overcoming Limitations of Lightweight Models for Unmanned Military Image Classification AI Model using Knowledge Distillation

Ja-Hoon Jeong¹, Yoon-Ho Song¹, In-Uk Kang², Jun-Yeol Ryu^{3*}

¹Department of Mechanical and Systems Engineering, Korea Military Academy

²Korea Research Institute for defense Technology Planning and Advancement

³Department of Industrial Engineering, Seoul National University

요약 군사 목적의 무인 체계에 활용되는 AI 모델을 개발함에 있어 군의 특수한 운용 환경을 고려하는 것은 중요하다. 예를 들면 사용 가능한 배터리 전력이 제한되어 있거나 최전선에서 정찰용으로 운용되는 무인체계의 경우 파괴 위험이 높아 고비용 고성능 칩을 사용하기 제한된다는 것이다. 이번 연구에서는 두 가지 문제 상황을 고려하여 AI 모델의 이미지 분류 성능향상 방법을 연구하였다. 첫 번째 문제는 무인 체계의 전력, 비용 등의 제약으로 인해 고용량 고성능 모델을 활용하기 제한된다는 것이다. 두 번째 문제는 군 AI 모델의 성능 보장을 위한 충분한 학습 데이터가 부족하다는 것이다. 이 문제에 대응하기 위해 본 연구에서는 지식증류기법을 제안한다. Teacher 모델로는 연산량이 많지만 성능이 우수한 EfficientNetB4를, Student 모델로는 SqueezeNet, ShuffleNetV2, MobileNetV3 small을 선정하였다. 지식 증류를 통해 정확도가 높은 Teacher 모델의 지식이 전달되어 효과적으로 Student 모델을 보강하여 제약이 있는 상황에서도 모델의 분류 성능이 향상되었다. 이와 같은 결과는 학습 데이터가 부족한 상황에서 탑재형 AI 모델에 적용되는 경량화 모델의 성능 부족 문제를 보완하여 군사적 활용성을 높일 수 있을 것으로 기대한다.

Abstract Developing AI models for military unmanned systems requires consideration of the unique operational environment. Constraints like limited battery power and the high risk of destruction at the frontline necessitate restrictions on using costly, high-performance chips. In this study, we explored methods to enhance image classification performance of AI models under two key challenges. Firstly, constraints such as power and cost limit the utilization of high-capacity, high-performance models in unmanned systems. Secondly, there's a shortage of sufficient training data to ensure the performance of military AI models. To address these issues, we propose knowledge distillation. We selected EfficientNetB4 as the Teacher model, known for its superior performance despite high computational complexity, and SqueezeNet, ShuffleNetV2, and MobileNetV3 small as Student models. Through knowledge distillation, the high-accuracy knowledge of the Teacher model effectively enhanced the Student models, improving classification performance even under constraints. Such results are expected to enhance military utility by addressing the performance limitations of lightweight models applied to on device AI model in scenarios with limited training data.

Keywords : Knowledge Distillation, Image Classification, Military Unmanned Systems, EfficientNetB4, SqueezeNet, MobileNetV3 small, ShuffleNetV2, LightWeight Model, On Device AI

이 논문은 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임. (No. RS-2023-00222172)

*Corresponding Author : Jun-Yeol Ryu(Department of Industrial Engineering, Seoul National University.)

email: barampool@gmail.com

Received February 21, 2024

Revised March 7, 2024

Accepted March 8, 2024

Published March 31, 2024

1. 서론

국가통계포털, 국방 백서, 국방혁신 4.0 등에서 명시된 바와 같이 인구절벽으로 인해 입영자원이 감소하고 있으며, 복무기간의 단축으로 인해 전투 임무 수행, 정보 분석 등을 수행할 수 있는 숙달된 병력의 수가 감소하고 있다[1]. 이 문제를 해결하기 위해 전 세계적으로 무인 무기체계(Autonomous Weapon Systems)가 활발히 개발되고 있으며 현대 전장에서 그 효용이 증명되었다. 예를 들면 최근 우크라이나-러시아 전쟁에서는 다양한 공중 무인체계를 사용하여 적지 중심 정찰, 실시간 전장 가시화, 중요 시설 파괴 및 인명 살상 등의 군사적 성과를 달성하였다. 이스라엘-하마스 간의 전쟁에서는 지상 무인체계를 활용하여 하마스 땅굴 탐색 등의 임무를 수행하여 인명 손실을 줄이고 효과적인 작전 임무를 수행하였다. 무인체계는 사람을 대신하여 임무를 수행하기 때문에 다양한 센서 등을 통해 획득한 정보가 적인지 아군인지를 분류할 수 있어야 할 뿐만 아니라 적 무기체계의 종류를 파악할 수 있어야 한다. 이를 바탕으로 표적 지시, 공격 무기체계의 선택 등 군사적 임무를 수행할 수 있다.

그러나 전장의 특성으로 인해 무인체계에 적용되는 AI 모델이 직면하는 대표적인 문제가 있다. 첫 번째는 학습 데이터가 부족하다는 것이다. AI 모델의 성능을 향상하기 위해서는 충분한 데이터가 필요하다. 그러나 평시적 무기체계에 대한 정보는 제한적이다[2,3]. 적이 수출하는 무기체계의 표준형 데이터를 사용하거나 열병식 등에 공개되는 데이터를 사용하는 등 수집할 수 있는 데이터가 제한적이다. 더욱이 전시에는 다양한 형태 및 종류의 부가 무장, 위장 패턴 등 평소 학습한 데이터와 실제 무기체계 간의 차이가 존재한다. 그러므로 초기 전투에서 식별한 데이터를 사용하여 AI 모델을 학습시켜야 한다. 두 번째는 무인 체계에 사용되는 AI 모델은 탑재형(On Device)이라는 점이다. 배터리와 같이 장착된 체계 사양 등에 따라 제한적인 성능을 갖기 때문에 전력, 내장 컴퓨터의 성능, 비용 등의 제약이 존재하여 일반적으로 사용하는 고용량, 고성능 모델을 사용하기 어렵다. 예를 들면 전방 감시 무인체계와 같이 운용 목적상 빈번히 파괴될 것으로 예상되는 무기체계의 경우 고비용의 부품을 사용하기 제한된다. 따라서 이 경우 탑재형 AI 모델의 제한사항을 고려하여 경량화된 모델을 사용해야 한다.

경량화 모델의 경우 모델의 구조가 단순하고 학습할 수 있는 파라미터 및 하이퍼 파라미터가 부족하여 고용

량 모델보다 성능이 낮을 가능성을 내포하고 있다. AI 모델을 경량화하는 방법은 대표적으로 구조적 축소, 가지치기와 지식증류(Knowledge Distillation) 기법이 있다. 구조적 축소 기법과 가지치기 방법은 인위적인 선택으로 인해 과적합 문제를 야기할 수 있지만, 이에 반해 지식증류 기법은 데이터 기반의 방법으로 과적합 위험을 줄이고 모델의 일반화 능력을 향상할 수 있는 방법이다. 지식증류 기법은 우수한 성능을 갖는 비교적 고성능 모델에서 학습된 지식을 비교적 성능이 떨어지는 모델로 지식을 전달함으로써 성능을 향상하는 방법이다. 지식을 전달하는 고성능 모델은 Teacher 모델이라 불리고, 지식을 전달받는 비교적 저성능 모델은 student 모델이라 불린다. 이때, 통상적으로 모델의 규모가 클수록 성능이 좋으므로 Teacher 모델은 큰 규모, Student 모델은 비교적 작은 규모이다. 이를 통해 Student 모델의 성능 부족 문제를 효과적으로 보완할 수 있다.

본 논문에서는 지식증류 기법을 군사 분야에 적용하여 데이터가 부족한 상황에서도 AI 모델을 경량화하면서 동시에 성능을 향상할 수 있다는 것을 연구하였다.

본 논문은 총 4장으로 구성되어 있다. 2장에서는 관련 연구를 고찰하고 연구 방법과 이번 실험에 활용한 모델에 대해 설명한다. 3장에서는 지식증류 방안을 사용한 연구 결과를 분석한다. 마지막 4장에서는 결론 및 향후 연구 방향에 대해 제시한다.

2. 본론

2.1 관련 연구

지식 증류는 2015년 9월 "Distilling the Knowledge in a Neural Network" 논문에서 처음으로 공개되었다. G. Hinton 등 3명은, 모델의 성능을 쉽게 올리는 방법의 하나인 앙상블(Ensemble)의 경우 높은 Computing Cost가 필요하기에 일반적인 상황에 제한됨을 인식하고, 앙상블을 통해 얻을 수 있는 일반화(Generalization) 능력을 더 작은 규모의 모델에 전달하는 방법을 착안하였는데, 이것이 지식 증류이다. 이는, 지식을 전달해 주는 크고 복잡하지만, 정확도가 높은 모델을 Teacher 모델로 정하고, 가볍지만 복잡하지 않은 모델을 Student 모델로 정하여 Teacher 모델의 학습 결과를 Student 모델로 증류하여 모델의 성능을 향상하는 방법이다. 다시 말해 지식을 받아 성능을 향상하는 방법을 말한다[4]. 이 방법론은 다음 페이지 Fig. 1에 설명돼 있다.

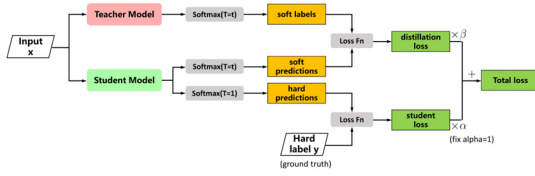


Fig. 1. The Framework of Knowledge Distillation (Classification)

이는 분류(Classification) 문제에서 동일한 입력 데이터(Input)를 활용해 Teacher 모델과 Student 모델을 학습시킨 뒤, Teacher 모델이 도출하는 Softmax 결과 (Input 값이 각 클래스에 속할 확률)에 Temperature Parameter(이하 T)를 활용하여 Soft labeling의 결과를 만든다. 동일하게 Student 모델이 도출하는 Softmax 결과에도 T를 활용하여 Soft Prediction을 만들어 이를 활용하여 Distillation Loss(중류 손실)를 정의한다. 또한, Student 모델에서는 Input을 통해 도출된 Hard Predictions(Softmax 결과 중 가장 높은 확률을 가진 클래스를 1, 나머지를 0으로 처리)과 Ground Truth를 비교하는 Student Loss를 정의한다. 이렇게 정의된 두 손실을 바탕으로 α, β 를 조절 변수로 하여 Total Loss를 정의한다. 결국 이 모델은 이렇게 정의된 Total Loss를 최소화하는 방향으로 학습이 진행된다. 이때, T는 도출된 Softmax 결과인 지식(Knowledge)의 크기에 영향을 미치는 변수이다. 아래 Eq. (1)에서 정의된 총 손실함수(Total) 식을 보면, T는 Student 모델의 네트워크와 Teacher 모델의 네트워크에서 도출된 Output logits (softmax 값)을 나누어주며, 해당 Cross Entropy Loss 함수에 T의 제곱이 곱해짐을 알 수 있다. 특히, T는 Teacher 모델과 Student 모델의 Output logits를 나누어 각 클래스에 속할 확률 분포를 극단적으로 만들지 않는 역할을 하며 이를 Temperature로 표현하여 지식을 중류(Distillation)한다는 표현을 사용했다. 결정 변수는 오른쪽 Table 1에 설명된 것과 같다.

$$Total Loss = (1 - \alpha)L_{CE}(\sigma(Z_s), \hat{y}) + 2\alpha T^2 L_{CE}(\sigma(\frac{Z_s}{T}), \sigma(\frac{Z_t}{T})) \quad (1)$$

이와 같은 지식중류 방법은 다양한 분야에서 활발히 활용되고 있다. Shiqi Chen et al.과 Zhen Wang et al.은 위성에서 활용되는 SAR의 영상의 분석을 위해 가지치기(Pruning) 및 지식중류 방법을 활용하여 경량화된 네트워크를 만들었다[5,6]. Guyang Yu는 UAV(Unmanned

Table 1. Notation

Notation	Description
$L_{CE}()$	Cross entropy loss
$\sigma()$	Softmax
Z_s	Output logits of Student network
Z_t	Output logits of Teacher network
\hat{y}	Ground truth(one-hot)
α	Balancing Parameter
T	Temperature hyperparameter

Aerial Vehicle)에 활용되는 영상 분석모델의 성능향상을 위해 GAN(Generative Adversarial Network)을 활용하여 학습모델의 성능을 향상한 다음 이를 지식중류하여 사용모델의 성능을 향상하는 지식중류 방법을 제안하였다[7]. Haoyu Lou et al.은 UAV를 활용한 배송서비스에서 정확한 배송지 위치 인식을 위해 대용량 클라우드 서버 기반의 Teacher 모델이 학습한 성능 높은 지식을 UAV에 탑재된 경량화된 모델에 전달하는 지식중류 방법을 활용하였다[8]. Houzhang Fang et al.은 UAV에 수집되는 적외선 영상을 여러 크기의 수준으로 학습하여 습득된 지식을 전달하여 물체의 위치를 탐지하는 모델을 제안하였다[9].

이와 같이 많은 논문에서 지식중류 기법을 활용하여 탑재용으로 설계된 경량화 모델의 성능을 향상하거나 복잡한 학습 과정의 지식을 효과적으로 압축하여 전달하는 방식을 활용하고 있다[10]. 하지만, 의도적으로 위장된 데이터가 많은 군사적인 데이터를 활용해 직접적으로 지식중류 기법을 군사 분야에 적용한 연구는 부족한 상황이다. 본 연구에서는 수집된 군사 이미지 데이터를 기반으로, 파라미터 수가 많고 성능이 높은 Teacher 모델을 선정 및 비교하고, Teacher 모델에 비해 파라미터수가 적지만 성능이 준수한 경량화 모델 중 대표적인 모델을 선정하여, 부족한 데이터라는 한계점과 가벼운 모델로부터 발생하는 제한된 성능이라는 한계점을 극복하는 방법론을 제안하고자 한다.

2.2 연구 방법

2.2.1 데이터 처리 및 모델 선정

이번 연구에서 활용한 데이터는 러시아-우크라이나 전쟁 등 적성 장비 중 수집할 수 있는 기갑 및 기계화 차량 6종류(BMP, BTR, T-54, T-55, T-62, T-72)의 이미지 데이터이다. 이와 같이 구성한 이유는, NATO 이외 군의 주력 전차 및 장갑차 전력이며, 서로 다른 바퀴의

형태, 방호력, 무장을 갖고 있어 소부대 교전 수준에서 상이한 대응 방법을 구사해야 하기 때문이다. 기종별 이미지 데이터의 구성 및 모습은 아래 Table 2와 같다.

Table 2. Classes of the Data Set

Image		
Class	BMP	BTR
# of Data	101	72
Image		
Class	T-54	T-55
# of Data	81	76
Image		
Class	T-62	T-72
# of Data	86	94

위 그림에서 보는 것처럼 분석에 활용된 데이터는 클래스 별로 부족하며, 클래스 간의 불균형이 있는 Long-tailed distribution을 보인다. 이는 외부 공개가 제한되는 군사적 데이터의 특성을 잘 나타내고 있다.

이러한 상황에서도 모델이 과적합(Overfitting) 되지 않고, 준수한 일반화 능력을 얻기 위해선 메타학습(Meta Learning), 데이터 증강 등의 방법이 있다. 이 중 본 연구에서는 전이학습을 기반으로 한 지식증류 방안을 선택하였다. 이는, 대규모의 이미지 데이터로 사전 학습된 네트워크들이 가지고 있는 특성 추출(Feature Extraction) 능력을, 규모가 작은 탑재형 AI에 효과적으로 전달하여 모델의 성능을 향상시킬 수 있는 방법이기 때문이다. 이는, 전장의 불확실성으로 인해 서버에 연결되지 않고도 효과적으로 임무를 수행해야 하는 탑재형 AI에 매우 적합하기 때문이다. 또한, 모델의 데이터에 대한 강건성을 위해 이미지 데이터에 임의로 변형(Cropping, Shifting, Rotating)을 가했다.

2.2.2 Teacher 및 Student 모델 선정

먼저, Teacher 모델을 선정하였다. 데이터가 부족한 상황임을 감안해 귀납 편향(Inductive bias)이 부족한

ViT기반 모델은 제외하고 CNN에 기반한 좋은 성능을 내는 모델을 선정하였다. 모델 선정을 위해 ImageNet으로 사전 학습된 결과를 가진 모델 중에서 top-5 error가 인간의 top-error인 5.1%보다 낮은 모델을 후보군으로 정했다. 이 중 군용 항공기의 실시간 이미지 분류에서 효과를 보여준 ResNet 50과 ResNet50 보다 더 많은 Residual block으로 인해 더 깊게 모델의 추상적인 정보까지 학습할 수 있는 ResNet 152, 그리고 Neural Architecture Search(NAS)를 사용하여 최적화된 아키텍처를 찾아낸 EfficientNet 시리즈 중 비교적 크고 높은 정확도를 나타내는 EfficientNetB4를 Teacher 모델의 후보로 선정하였다[11,12]. 이 세 모델을 활용해 Train 데이터로 학습 후 Validation 데이터로 검증한 결과는 다음 페이지 Table 3과 같다. 이때, Epochs는 100, Optimizer는 AdamW(Weight Decay=0.001)로 동일하게 설정하고 모두 ImageNet 데이터로 사전 학습된 가중치를 활용하여 검증하였다. 학습 및 검증 결과, ResNet 152가 가장 낮은 Validation Loss를 기록했으나 이와 유사한 Validation Loss를 기록하고 가장 높은 Test Accuracy를 보인 EfficientNetB4를 최적의 모델로 판단하여 Teacher 모델로 선정하였다.

Student 모델로 선정한 것은 1) 2016년 AlexNet의 1/50 파라미터 수로 유사한 성능을 보여준 가벼운 모델인 SqueezeNet, 2) 2018년 Group Convolution과 Shuffling을 결합한 Shuffle Unit을 활용하여 구축된 ShuffleNetV2, 3) MobileNetV2와 유사한 성능을 보이지만 더 적은 파라미터수로 인해 경량화 및 빠른 학습속도, 그리고 준수한 성능을 보이는 MobileNetV3 Small이다[13-16]. 각 모델 별 파라미터의 수는 SqueezeNet은 5 Millions, ShuffleNetV2는 2.2 Millions, MobileNetV3 Small은 2.5 Millions로 선정한 Teacher 모델보다 약 1/30 ~ 1/16 정도 크기의 파라미터 수를 가지고 있는 경량화 모델이다. 먼저, Student 모델의 성능을 확인하기 위해 학습 및 검증 데이터 세트를 활용해 학습을 시켰으며, 이 결과는 다음 페이지 Table 4와 같다. 이때, Squeeze Net과 MobileNetV3 small은 100 Epochs로 학습하였을 때 학습용 데이터에 대해 과적합(Overfitting)이 발생하여 가장 낮은 Validation Loss를 기록하는 최적의 모델을 찾았으나, ShuffleNetV2의 경우 학습 속도가 느려 300 Epochs로 최적의 모델을 찾을 수 있었다. 또한, Optimizer는 AdamW(Weight Decay = 0.001)로 동일하며, 모두 ImageNet 데이터로 사전 학습된 가중치를 활용하였다. Student 모델 별 최적의 모델로 테

Table 3. Classification Results of 3 CNN (Teacher Model)

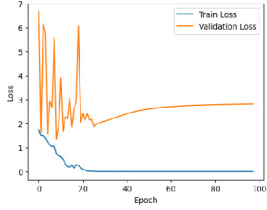
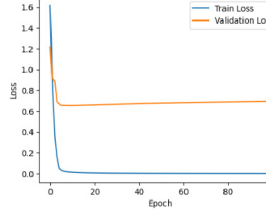
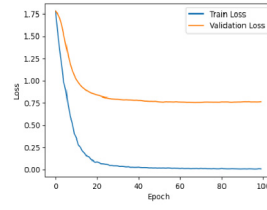
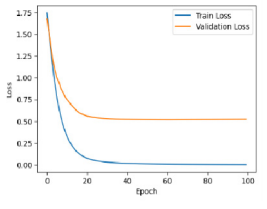
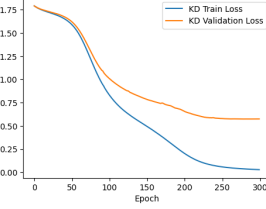
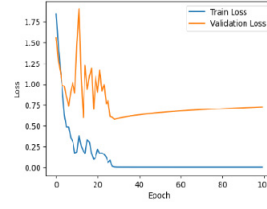
Model	ResNet-50	ResNet-152	EfficientNetB4
# of Parameters	25 Millions	60 Millions	66 Millions
Train & Validation Loss			
Test Accuracy (Best Model)	85%	80%	90.91%

Table 4. Classification Results of 3 CNN Models (Student Model)

Model	SqueezeNet	ShuffleNetV2	MobileNetV3 Small
# of Parameters	5 Millions	2.2 Millions	2.5 Millions
Train & Validation Loss			
Test Accuracy (Best Model)	65.15%	74.24%	80.30%

스트 데이터에 대해 평가하였고, SqueezeNet은 65.15%, ShuffleNetV2는 74.24%, MobileNetV3 small은 80.30%의 Test Accuracy를 보여주었다.

2.2.3 실험 환경

이번 연구의 실험 환경은 구글 코랩(Google Colaboratory) 환경을 활용하였으며, CPU는 Intel (R) Xeon(R) CPU @ 2.3GHz이고, GPU는 Nvidia Tesla T4 8GB이며, Memory는 16GB, 그리고 Single-Precision Performance는 5.5 TeraFLOPs이다.

3. 지식증류 방안 성능 평가

2장에서 설명한 바와 같이, Teacher 모델은 EfficientNetB4이며, Student 모델은 Squeezenet과

ShuffleNetV2, MobileNetV3 small이다. 아래 Table 5와 같이 Student 모델은 지식 증류 전 Teacher 모델과 Test Accuracy에서 10~25% 정도의 성능 차이를 보였다.

Table 5. Classification Results of 3 CNN Models (Students)

Model	Teacher	Student		
	EfficientNetB4	SqueezeNet	ShuffleNet V2	MobileNet V3 Small
Test Accuracy	90.91%	65.15%	74.24%	80.30%
Difference from Teacher	-	25.76%	16.67%	10.61%

Teacher 모델의 지식인 Softmax 결과가 더 정확한 예측을 한 것을 확인했으니, 다음으로 지식 증류를 통해 Student 모델의 분류 결과가 향상되는지를 확인하였다. 이번 연구에서는 지식 증류 간 파라미터 T를 1에서 6까지 총 6개의 범주로 나누어 변수를 조정하여 그 결과를 확인하였다. 이때 T를 6의 범주까지 적용한 것은 데이터의 총 클래스의 개수를 고려하여 적용했으며, Epochs는 각 100회로 제한하였다.

SqueezeNet에 T=1, 2, 3, 4, 5, 6을 적용한 지식 증류의 결과는 아래 Table 6과 같다.

Table 6. Test Accuracy after Applying Knowledge Distillation(Student : SqueezeNet)

T	1	2	3	4	5	6
Test Accuracy	84.85%	78.79%	75.76%	77.27%	80.30%	72.73%

SqueezeNet을 단독으로 학습했을 때 Test Accuracy가 65.15%인 것을 고려하면, 모든 T에 대해서 지식증류의 효과가 있음을 확인할 수 있다. 특히, T가 1인 경우 약 19%의 Test Accuracy의 상승을 확인할 수 있으며, 이 경우가 가장 지식이 잘 전달되었다고 할 수 있다.

ShuffleNet에 T=1, 2, 3, 4, 5, 6을 적용한 지식 증류의 결과는 아래 Table 7과 같다.

Table 7. Test Accuracy after Applying Knowledge Distillation(Student : ShuffleNetV2)

T	1	2	3	4	5	6
Test Accuracy	75.76%	89.39%	69.70%	72.73%	72.73%	74.24%

ShuffleNet을 단독으로 학습했을 때 Test Accuracy가 74.24%인 것을 고려하면, T가 1, 2인 경우 1~15%의 Test Accuracy의 향상이 있었지만, T가 3, 4, 5, 6인 경우 Test Accuracy가 같거나 더 떨어지는 경우를 확인할 수 있었다. 이때, T가 2일 때 가장 지식이 잘 전달되었으며 Teacher 모델의 Test Accuracy인 90.91%에 근접한 결과를 만들 수 있었다.

MobileNetV3 Small에 T=1, 2, 3, 4, 5, 6을 적용한 지식 증류의 결과는 오른쪽 Table 8과 같다.

Table 8. Test Accuracy after Applying Knowledge Distillation(Student : MobileNetV3 Small)

T	1	2	3	4	5	6
Test Accuracy	87.88%	83.33%	89.39%	86.36%	86.36%	86.36%

MobileNetV3 Small을 단독으로 학습했을 때 Test Accuracy가 80.30%인 것을 고려하면, 모든 T에 대해서 지식증류가 Test Accuracy의 성능향상을 도와줬음을 알 수 있다. 특히, T가 3일 때는 Teacher 모델의 Test Accuracy인 90.91%에 근접한 결과를 만들 수 있었다.

지식증류 방안을 적용하여 모델의 분류 성능 향상을 확인한 다음, 지식증류가 되지 않은 Student 모델이 오분류한 Test 데이터 이미지가 지식 증류의 과정을 통해 어떻게 제대로 분류되는지를 Grad CAM을 통해 확인하였다. 이때, Grad-CAM은, 활성화 맵을 구할 때 GAP(Global Average Pooling)를 사용할 수 있는 경우에만 연산이 가능하다는 한계를 가지고 있어 그 활용이 제한적인 CAM(Class Activation Map)의 단점을 보완하여 나온 개념이다. Grad-CAM(Gradient weighted Class Activation Map)은 모델의 예측 클래스에 대한 그레디언트를 이용하여 채널별 특성 맵의 중요도를 계산하고, 0보다 큰 그레디언트가 높은 부분이 모델 예측에 미치는 영향이 큰 것으로 여겨 가중 평균을 구하고, 이를 특성 맵에 반영하여 클래스 활성화 맵을 구한다.

아래 Table 9는 실제 T=55 이미지를 Grad CAM을 활용해 Student 모델, Teacher 모델, Student with KD(Knowledge Distillation) 모델의 분류 결과를 Grad CAM을 활용해 시각화한 것이다. 마찬가지로 다음 페이지 Table 10은 BMP를, Table 11은 BTR을 시각화한 것이다.

Table 9. Classification results of the knowledge distillation model visualized with Grad-CAM(T=55)


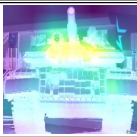

Image Class	Student	Teacher	Student with KD
T-55			
Model	SqueezeNet	EfficientNetB4	KD (T=1)
Prediction	BTR	T-55	T-55

Table 10. Classification results of the knowledge distillation model visualized with Grad-CAM(BMP)







Image Class	Student	Teacher	Student with KD
BMP			
Model	ShuffleNetV2	EfficientNetB4	KD (T=2)
Prediction	T-54	BMP	BMP

Table 11. Classification results of the knowledge distillation model visualized with Grad-CAM(BTR)

Image Class	Student	Teacher	Student with KD
BTR			
Model	MobileNetV3 Small	EfficientNetB4	KD (T=3)
Prediction	BMP	BTR	BTR

이때, 사용된 데이터는 Test 데이터 중 각 Student 모델은 잘못 분류하고 Teacher 모델은 제대로 분류한 데이터 중 일부를 활용하였으며, 분류된 지식이 전달된 Student 모델은 2.4.1.부터 2.4.3.에 정리된 각 모델 별 가장 Test Accuracy를 나타낸 T를 사용하였다. 아래 그림에서 보는 것처럼, Student 모델은 각 전차/장갑차가 위치한 픽셀이 아닌 주변 픽셀에 중점을 두고 분류하였고, Teacher 모델은 각 장비가 위치한 픽셀에 중점을 두고 분류하였다. 그리고, 지식 증류를 통해 Teacher 모델의 지식이 전달된 Student 모델의 결과를 보면, 기존 Student 모델 보다 전차/장갑차가 위치한 픽셀에 중점을 두어 무기체계의 특징적 부분에 집중하여 이미지를 분류하고 이를 통해 해당 모델은 실제 클래스를 더욱 정확하게 예측할 수 있었다.

4. 결론 및 향후 연구 방향

이번 연구는 유·무인 및 무인화 되어가는 현대전의 추세에 따라 개발되는 이미지 분류 AI의 발전 방향 중 하나

인, 탑재형 AI에 사용되는 가벼운 모델의 성능을 향상하는 방안을 제시한다. 이와 같은 가벼운 모델의 한계(복잡한 모델을 만들지 못해 고차원의 정보 추출 제한)와 더불어 군사 데이터의 특성상 공개된 자료가 적다는 "데이터의 희소성"까지 더해져, 전선에서 사용되는 자율 무기체계의 성능 개선에 한계가 있다. 이를 위해 지식 증류라는 기법을 통해 무겁고 고차원의 정보를 추출할 수 있는 Teacher 모델의 지식을 증류하여 가벼운 Student 모델에 전달하였다. 실험 결과, 각 Student 모델은 최대 19%에 달하는 분류 성능의 향상을 확인할 수 있었다. 이에 대한 결과를 Grad CAM이라고 하는 XAI(eXplainable AI) 기법을 활용해 CNN 기반 모델의 분류 결과를 시각적으로 확인할 수 있었고, Student 모델은 Teacher 모델의 지도에 따라 그 중점이 분류 클래스의 주변 이미지로부터 분류 클래스의 중심 이미지로 점차 이동하였음을 알 수 있었다. 이러한 결과는 운용환경과 데이터 제약을 고려한 군사용 AI 모델의 향상에 활용할 수 있을 것으로 기대한다.

이번 연구의 한계점으로는 데이터 수의 부족으로 인한 결과의 불안정성이다. 먼저, 공개된 데이터의 수가 충분치 못한 군사 데이터의 특성상 이러한 데이터 부족 문제는 늘 수반되는 문제이며, 부족한 데이터로 학습한 것에 대한 결과의 안정성이 떨어진다는 부분을 한계점으로 들 수 있다. 하지만, 이러한 상황에서도 인위적인 데이터 증강 방법이 아닌 Teacher 모델과 Student 모델의 네트워크를 활용한 지식증류 기법을 통해, 모델의 구조적인 관점에서 이를 해결할 수 있었다. 향후 연구에서는 보다 다양한 데이터를 수집하고, 2D 이미지를 3D로 만드는 생성형 모델을 활용해 보다 다양한 관점의 데이터를 수집하여 보다 신뢰성있는 네트워크를 구성하여 믿을 수 있는 자율형 무기체계 구축에 관해 알아보려고 한다.

References

- [1] Ministry of Defense(MnD), Republic of Korea "Defense White Paper 2022", p.107-114, MnD, Republic of Korea., 2023.
- [2] Svenmarck, P., Luotsinen, L., Nilsson, M., & Schubert, J., "Possibilities and challenges for artificial intelligence in military applications", *In Proceedings of the NATO Big Data and Artificial Intelligence for Military Decision Making Specialists Meeting*, pp. 1-16, May 2018.
- [3] Rettore, P. H., Zißner, P., Alkhawaiter, M., Zou, C., & Sevenich, P. "Military data space: challenges, opportunities,

- and use cases", *IEEE Communications Magazine*, Vol. 62, Issue 1, pp.70-76, Jan 2024.
DOI: <https://doi.org/10.1109/MCOM.001.2300396>
- [4] G. Hinton, O. Vinyals, J. Dean, "Distilling the Knowledge in a Nueral Network", *arXiv preprint arXiv: 1503.02531*, Mar 2015.
DOI: <https://doi.org/10.48550/arXiv.1503.02531>
- [5] Chen, S., Zhan, R., Wang, W., & Zhang, J. "Learning slimming SAR ship object detector through network pruning and knowledge distillation", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 14, pp.1267-1282, 2020.
DOI: <https://doi.org/10.1109/JSTARS.2020.3041783>
- [6] Z. Wang, L. Du and Y. Li, "Boosting Lightweight CNNs Through Network Pruning and Knowledge Distillation for SAR Target Recognition", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 14, pp. 8386-8397, 2021.
DOI: <https://doi.org/10.1109/JSTARS.2021.3104267>
- [7] G. Yu, "Data-Free Knowledge Distillation for Privacy-Preserving Efficient UAV Networks", *2022 6th International Conference on Robotics and Automation Sciences (ICRAS)*, Wuhan, China, pp. 52-56, 2022.
DOI: <https://doi.org/10.1109/ICRAS55217.2022.9842039>
- [8] Luo, H., Chen, T., Li, X., Li, S., Zhang, C., Zhao, G., & Liu, X., "KeepEdge: A knowledge distillation empowered edge intelligence framework for visual assisted positioning in UAV delivery", *IEEE Transactions on Mobile Computing*, Vol. 22, Issue 8, pp.4729-4741, 2022.
DOI: <https://doi.org/10.1109/TMC.2022.3157957>
- [9] H. Fang, Z. Liao, L. Wang, Q. Li, Y. Chang, L. Yan, "DANet: Multi-scale UAV Target Detection with Dynamic Feature Perception and Scale-aware Knowledge Distillation", *Proceedings of the 31st ACM International Conference on Multimedia*, Vol. 31, pp. 2121-2130, 2023.
DOI: <https://doi.org/10.1145/3581783.3612146>
- [10] Gou, J., Yu, B., Maybank, S. J., & Tao, D., "Knowledge distillation: A survey", *International Journal of Computer Vision*, Vol. 129, pp.1789-1819, 2021.
DOI: <https://doi.org/10.48550/arXiv.2006.05525>
- [11] C. V. Sudhakar, D. Limbakar, K. Charan, P. Sah, Rauniyar, K. Mahathi, "Real Time Web-base System to Detect Military Aircraft Using RESNET-50 Algorithm", *Electrical and Automation Engineering*, Vol. 2(1), pp.89-96, 2023.
DOI: <http://dx.doi.org/10.46632/ae/2/1/13>
- [11] Tan, M., & Le, Q., "Efficientnet: Rethinking model scaling for convolutional neural networks" *In International conference on machine learning Proceeding of Machine Learning Research*, pp. 6105-6114, May 2019.
DOI: <https://doi.org/10.48550/arXiv.1905.11946>
- [12] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization", *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp.618-626, 2017.
DOI: <https://doi.org/10.1109/ICCV.2017.74>
- [13] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet classification with deep convolutional neural networks", *Proceedings of the 25th International Conference on Neural Information Processing Systems*, Vol. 1, pp.1097-1105, 2012
DOI: <https://doi.org/10.1145/3065386>
- [14] Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K., "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size", *arXiv preprint arXiv:1602.07360*, 2016
DOI: <https://doi.org/10.48550/arXiv.1602.07360>
- [15] Zhang, X., Zhou, X., Lin, M., & Sun, J., "Shufflenet: An extremely efficient convolutional neural network for mobile devices", *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6848-6856, 2018
DOI: <https://doi.org/10.48550/arXiv.1707.01083>
- [16] Koonce, Brett, and Brett Koonce. "MobileNetV3." *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*, p.245, Apress, 2021, pp.125-144.
DOI: <https://doi.org/10.1007/978-1-4842-6168-2>

정 자 훈(Ja-Hoon Jeong)

[정회원]



- 2021년 7월 : North Carolina State University Industrial & System Engineering (공학석사)
- 2021년 8월 ~ 2022년 10월 : 육군사관학교 기계·시스템공학과 조교수

- 2022년 11월 ~ 2023년 10월 : 육군대학
- 2023년 11월 ~ 현재 : 육군사관학교 기계·시스템공학과 조교수

<관심분야>

무기체계, Computer Vision, Image Segmentation, Domain Adaptation, Time-series data analysis

송 윤 호(Youho Song)

[정회원]



- 2015년 2월 : 육군사관학교 전자공학과 (공학사, 군사학사)
- 2023년 1월 : 국방대학교 관리대학원 국방과학학과 (공학석사)
- 2023년 1월 ~ 현재 : 육군사관학교 기계 · 시스템공학과 강사

<관심분야>

무기체계, 국방M&S, 표적 탐지 및 추적, RCS

강 인 욱(In-Uk Kang)

[정회원]



- 2017년 2월 : 육군사관학교 무기시스템공학 (공학사)
- 2022년 8월 : 텍사스 A&M대학교 전기공학과 (공학석사)
- 2023년 3월 ~ 현재 : 국방기술진흥연구소 연구원

<관심분야>

전자전, 국방기술

류 준 열(Jun-Yeol Ryu)

[정회원]



- 2020년 2월 : 서울대학교 산업공학과 (공학석사)
- 2020년 1월 ~ 2023년 2월 : 육군사관학교 기계 · 시스템공학과 강사
- 2023년 2월 ~ 2024년 2월 : 국방기술진흥연구소 연구원
- 2024년 2월 ~ 현재 : 서울대학교 산업공학과 박사과정

<관심분야>

무기체계, 데이터 분석, Domain Adaptation, Engineering Design