

적록색맹 모사 영상 데이터를 이용한 딥러닝 기반의 위장군인 객체 인식 성능 향상

최 근 하^{*,1)}

¹⁾ 육군교육사령부 인공지능연구발전처

Performance Improvement of a Deep Learning-based Object Recognition using Imitated Red-green Color Blindness of Camouflaged Soldier Images

Keun Ha Choi^{*,1)}

¹⁾ *Artificial Intelligence Research & Development Center, Army Training and Doctrine Command, Republic of Korea*

(Received 20 January 2020 / Revised 25 March 2020 / Accepted 30 March 2020)

Abstract

The camouflage pattern was difficult to distinguish from the surrounding background, so it was difficult to classify the object and the background image when the color image is used as the training data of deep-learning. In this paper, we proposed a red-green color blindness image transformation method using the principle that people of red-green blindness distinguish green color better than ordinary people. Experimental results show that the camouflage soldier's recognition performance improved by proposed a deep learning model of the ensemble technique using the imitated red-green-blind image data and the original color image data.

Key Words : Deep Learning(심층 신경망학습), Camouflaged Soldier Recognition(위장군인인식), Ensemble Learning(앙상블 학습), Red-green Color Blindness(적록색맹)

1. 서론

객체인식이란 컴퓨터 비전 기술 중 하나로 영상 또는 비디오 영상으로부터 획득된 영상 상의 객체(물체)를 식별하는 기술로 최근에는 딥러닝으로 대변되는

머신러닝 기술의 비약적 발전으로 탁월한 성능을 보여주고 있다. 민수분야에서 객체인식 기술은 산업현장, 무인자동차, 의료 분야 등에서 다양하게 활용되고 있으며 향후 비약적 발전 가능성도 매우 높다고 할 수 있다. 하지만 본 저자가 [1]의 논문에서 기술한 바와 같이 민간분야에 광범할 만한 성능향상에도 불구하고 군사적 적용에서는 군사환경의 어려움(위장, 양질의 데이터 부족, 높은 신뢰성 필요)으로 인해 객체

* Corresponding author, E-mail: choiha99@kaist.ac.kr
Copyright © The Korea Institute of Military Science and Technology

인식 적용의 신뢰성을 높이는 쉽지 않다. 이는 군사 보안 측면에서 학습용 데이터의 양적 확보의 어려움도 큰 문제이나, 기존에 확보된 학습용 데이터 역시 인식하려는 객체가 주변 환경과 구분이 어려워 딥러닝을 통해 학습을 시키더라도 인식 성능을 민간 분야만큼 높이는 쉽지 않다는 점이다. 특히 위장패턴은 주변 수풀 환경(배경)과의 구분이 어려워 해당 장면을 찍은 컬러 영상을 학습용 데이터로 사용할 경우 배경 영상과 객체를 분류가 어렵다^[1].

본 논문에서는 이러한 점에 착안하여 영상의 배경과 구분이 어려운 위장패턴을 가진 객체의 인식률을 높이기 위한 학습용 영상데이터 생성 방법과 딥러닝 기반의 앙상블 학습 방법을 제안하고자 한다. 2장에서는 기존 객체인식 선행연구들을 정리하였으며, 3장에서는 위장 복장을 착용한 군인의 객체인식 실험결과를 제시하였고, 4장에서는 적록색맹 모사 학습용 영상데이터 생성 방법을 5장에서는 딥러닝 기반의 앙상블 학습 방법과 실험 결과를 기술하였다.

2. 선행연구(Previous Researches)

객체인식(Object recognition)과 객체탐지(Object detection)는 유사한 객체식별기술이지만 그 실행하는 방식에서는 차이점이 있다. 객체탐지가 영상내에서 객체의 인스턴스를 찾아내는 것으로 영상에서 객체를 식별하는 것 뿐 아니라 위치까지 탐지하는 것을 포함한다면 반면 객체인식은 영상에서 객체를 식별하는 분류 문제를 다룬다. 과거 객체인식 연구는 SIFT(Scale Invariant Feature Transform), SURF(Speed-Up Robust Features), Haar(Histogram of Oriented Gradients) 등과 같이 영상내 특정 객체가 가지는 특징점(Features)들을 검출하여 객체를 식별하는 이미지 프로세싱 방식을 주로 사용하였다.

최근에는 합성신경망(CNN: Convolutional Neural Network)을 이용한 AlexNet^[2], ZFNet^[3], VGG^[4], ResNet^[5], GoogLeNet^[6], DenseNet^[7]과 앙상블(Ensemble) 학습^[8] 등 다양한 딥러닝 모델이 등장하여 매우 뛰어난 인식률을 보여주고 있다. 이 중 AlexNet^[9]은 빠른 연산처리를 위해 2개의 GPU(Graphics Processing Unit)를 사용하여 8개 층의 CNN 구조로 이전 보다 더 깊은 신경망 구조를 제안하였으며, 활성화함수(Activation function)로 ReLU(Rectified Linear Unit)를 사용하여 기존보다 6배 이상

학습속도를 향상시켰다. 또한 전연결층(Fully-connected layer)에 Dropout 방법을 적용하여 높은 인식률의 향상을 보였다. AlexNet 이후 깊은 구조의 CNN 모델들이 등장하였으나, 이런 모델들이 어떻게 인식성능을 향상시키는지에 대한 이해가 부족하였다. ZFNet은 이런 물음에 답을 주었는데 CNN 내부의 각 층을 시각화하여 낮은 층에서는 선이나 모양 등과 같은 단순 특징들이 맵핑되고, 깊은 층에서는 객체의 형상과 유사한 특징들이 맵핑됨을 보였고, 합성곱 각 층에서 시각화된 정보를 이용하여 AlexNet 초기층을 수정한 ZFNet 방법을 제안하여 보다 높은 분류 정확도를 보였다. VGG는 매우 간단한 CNN구조로 우수한 성능을 보여준 모델로 합성곱 필터 크기를 3*3으로 설정하여 반복적으로 사용하는 것이 큰 필터를 사용하는 것 보다 효과적임을 보여주었다. GoogLeNet은 여러 크기의 합성곱층과 풀링층(Pooling)이 병렬적으로 수행되고 그 결과를 병합하는 방식의 Inception 모듈을 제안하여 다양한 특징을 추출하였다. 이런 구조는 여러 크기의 다양한 합성곱이 발생하므로 많은 연산량이 필요하게 되는데 이를 해결하기 위해 크기가 1*1인 합성곱을 사용하여 차원을 감소시키는 방법을 사용하였다. GoogLeNet은 9개의 Inception 모듈로 구성된 모델로 이후 Inception V2, V3, V4를 발표하며 지속 개선시켰으며, GoogLeNet에 ResNet을 병합한 Inception ResNet을 개발하였다. ResNet은 신경망의 깊이가 증가될 수록 정확도가 높아지지 않고 오히려 떨어지는 문제를 해결하기 위해 잔차학습(Residual learning) 방법을 제안하였다. ResNet은 잔차학습 개념을 VGG에 적용하여 모델을 개발하였다. DenseNet은 ResNet과 같은 모델이 특정층의 값이 최종 출력 값에 기여하는 정도가 적거나 없는 문제를 해결하여 성능을 향상시켰다. DenseNet은 모든 층이 ResNet의 잔차학습층에 연결되어 더 빠른 학습속도와 성능향상을 보여주었다.

최근 군사용용 분야에서도 딥러닝을 활용한 객체인식에 대한 연구가 진행되고 있으나 데이터의 부족과 군사적 환경의 어려움으로 아직 민간분야의 객체인식 성능에는 못 미치고 있는 것이 현실이다. 군사 분야에서는 현재 국방과학연구소를 중심으로 군사용 학습데이터의 부족문제를 해소하고자 학습데이터 증강에 대한 연구가 활발하게 진행되고 있는 중이다. Yang^[16,17]은 부족한 군사용 데이터의 보완을 위해 게임엔진 기반으로 생성한 딥러닝 기반 물체탐지용 가상 학습데이터의 효용성 평가실험을 수행하여 약 10 %의 성능

을 향상시킨 연구를 수행하였다. 또한, 비지도학습 대립훈련 기반 Image to Image 변환을 통해 가시광 영상과 적외선 영상 간 교차로 영상을 변환하는 연구를 수행하였고, 변환을 통해 합성된 새로운 영상들은 인공신경망 및 머신러닝 학습을 위한 데이터 증강의 활용 가능성을 보여주었다. Jin^[18]은 YOLO V2를 이용하여 19개의 무기체계를 학습한 후 정확도를 탐지 분석하는 연구를 통해 색깔이 비슷한 형상을 가진 무기체계의 경우 탐지성능이 낮음을 보여주었다. Oh^[19]는 그 래픽 기술로 생성한 가상의 가짜 영상 데이터를 기반한 딥러닝 학습 방법론의 경우에는 훈련단계와는 달리 실제 영상을 적용하는 추론단계에서 매우 낮은 예측성능을 보이는 단점 있는데, 전산실험 설계에 많이 사용되고 있는 FCCD 방법론을 적용하여 딥러닝 학습 데이터 구축을 위해 생성된 가상영상 데이터가 학습 효율에 미치는 영향성을 실험을 통해 확인하였다.

3. 위장군인의 객체인식

동물은 자신을 천적으로부터 보호하기 위해 주변의 색과 자신의 색을 비슷하게 만드는 방식으로 진화했다. 이 같은 보호색은 자신의 위치를 최대한 노출시키지 않기 위한 최고의 방법이자 최선의 방법이라 할 수 있다. 마찬가지로 군인도 전투에서 위장복장은 동물의 보호색과 같은 효과로 배경과의 구분을 어렵게 하여 적으로부터 보호하는 효과 때문에 모든 군인들은 이 위장복장을 모두 착용한다. 위장패턴이 주변 환경과 유사하다는 것은 영상처리 관점에서는 객체의 특징점(Features) 추출이 어려워 객체인식이 어렵다는 것을 의미한다고 할 수 있다. 과거 특징점 추출 방식의 객체인식 알고리즘에서도 이 같은 어려움은 발생될 수 있고, 최근 발전된 딥러닝 기법을 적용하더라도 객체인식의 어려움은 동일하게 발생될 것으로 생각된다. 실제로 구글의 Inception-v3 모델^[20]을 이용하여 유사한 환경 속의 위장군인 영상 데이터 100장을 대상으로 인식을 실험을 수행한 결과 Table 1과 같이 48 %의 낮은 인식률을 보였다^[1]. 인식성공 기준은 Inception-v3 모델을 통한 인식 결과 상위 5개 중 정답을 연상(사람, 군인, 방탄조끼, 소총, 군복 등)시킬 수 있는 결과가 1개라도 있다면 성공이라고 할 정도로 기준(Top-5 Error)을 낮추었지만 좋지 못한 성능을 보였다.



Fig. 1. Protective coloration and camouflaged soldier

Table 1. Classification accuracy of camouflaged soldier into similar background image using Inception-v3 model^[1]

Contents	Top-5 Accuracy
Test data (Images)	100
Classification accuracy	48 %

Fig. 2는 대표적으로 잘못 식별된 결과 영상을 보여주는데 이처럼 위장군인의 군복과 주변 숲 환경과의 색조 유사성 때문에 위장복장의 색 및 패턴이 다른 동물들과 유사하여 인식결과의 혼동이 발생한 것으로 생각된다. 또한 Inception-v3 모델에서 사용된 학습데이터(CIFAR-10) 중 위장복장을 착용한 군인 영상 데이터의 배경이 숲이 아닌 환경 영상이 다수 존재하여 인식을 저하가 발생된 것으로 보인다. 실제로 Fig. 2의 Result#1의 결과를 보면 학습데이터(CIFAR-10) 중에는 배경과 위장복장의 구분이 확연한 영상 데이터가 다수 학습에 활용되어 좋은 인식 결과를 보였으나, Result#2의 결과에서처럼 배경과 위장복장의 구분이 어려운 영상 데이터에서는 좋은 인식결과를 보이지 못한 것을 확인할 수 있다. 결국 인간의 육안으로는 구분이 가능 하지만, 기존의 딥러닝 객체인식을 이용한 방법으로는 낮은 인식률을 보이는 이유는 학습용 데이터의 질적 차이에 기인한 것으로 판단된다. 따라서, 본 논문에서는 유사한 배경 속 위장군인 인식을 높이기 위해서는 딥러닝 모델 뿐 아니라 학습용 데이터의 질적 향상도 요구된다고 생각되어 다음과 같은 학습용 영상 데이터 생성 방법을 제안하였다.


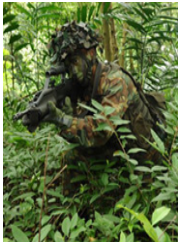
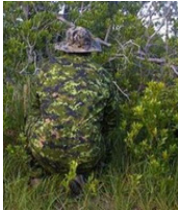
Image sample	Classification result(Inception-v3)
 <p>Result #1</p>	<p>Result #1 : Success</p> <p>bulletproof vest (score = 0.41013)</p> <p>military uniform (score = 0.31403)</p> <p>assault rifle, assault gun (score = 0.08510)</p> <p>cuirass (score = 0.01680)</p> <p>breastplate, aegis, egis (score = 0.01659)</p>
 <p>Result #2</p>	<p>Result #2 : Fail</p> <p>common iguana, iguana, Iguana iguana (score = 0.09010)</p> <p>frilled lizard, Chlamydosaurus kingi (score = 0.07735)</p> <p>tiger, Panthera tigris (score = 0.07458)</p> <p>orangutan, orang, orangutang, pongo pygmaeus (score = 0.06188)</p> <p>pedestal, plinth, footstall (score = 0.04550)</p>
 <p>Result #3</p>	<p>Result #3 : Fail</p> <p>barn spider, Araneus cavaticus (score = 0.16313)</p> <p>stone wall (score = 0.07785)</p> <p>spider web, spider's web (score = 0.07049)</p> <p>tailed frog, bell toad, ribbed toad, tailed toad, Ascaphus trui (score = 0.04834)</p> <p>megalith, megalithic structure (score = 0.04157)</p>

Fig. 2. Classification result examples of camouflaged soldier using Inception-v3

4. 적록색맹 모사 학습용 영상 데이터의 생성

적록색맹(Red-green color blindness)은 적색과 녹색의 식별 능력이 결여된 색각 이상을 말한다. 이런 색각 이상자가 일상생활에서는 분명 불리해 보이지만 꼭 불리한 것은 아니다. 일반인에게는 유사하게 보이는 색들이 색맹인 사람에게는 색이 확연하게 다르게 보인다는 원리를 이용해서 위장한 적을 잘 간파할 수도 있다. 일례로 2차 세계대전에서 독일군은 위장한 저격수를 찾는데 적록색맹인 군인을 활용하였다고 한다. 이는 적색맹이라서 녹색 위장을 더욱 쉽게 간파했다

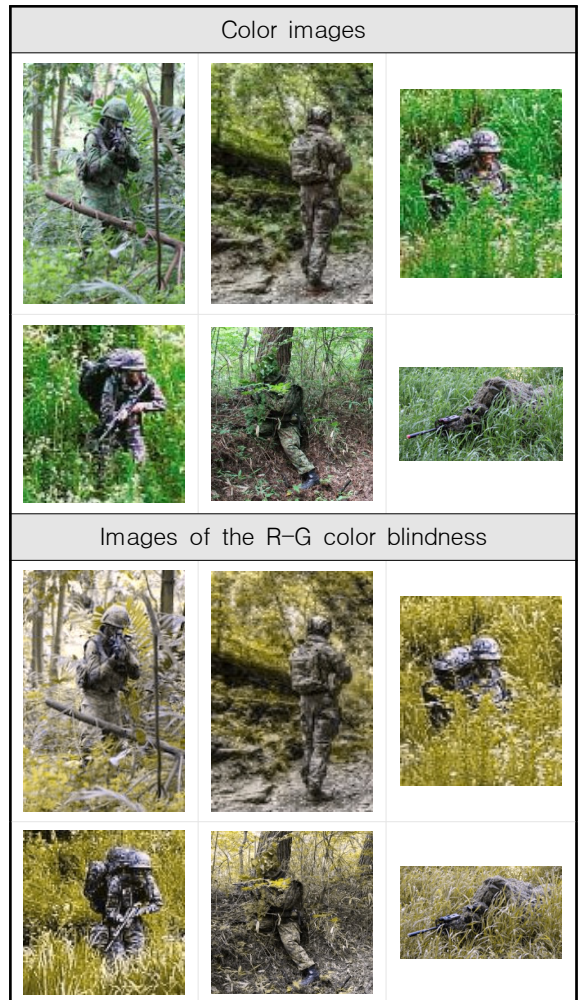


Fig. 3. Images of the red-green color blindness

고 보고 있으며, 적록색맹 유전자가 도태되지 않고 살아남은 이유가 바로 녹색의 미세한 차이에 더 민감해져서 풀숲에 숨은 사냥감을 더욱 쉽게 찾을 수 있었기 때문이라는 연구결과도 있다^[21].

본 논문에서는 이에 착안하여 컬러영상을 영상 전처리를 통해 적록색맹 영상을 모사하여 생성하였다. 적록색맹 영상 전처리 모델은 Hans Brettel^[22] 논문의 적록색맹 영상 변환 이론을 알고리즘으로 구현하였다. 영상의 컬러 공간(R,G,B)를 3차원의 LMS(Long-wavelength, Middle-wavelength, Short-wavelength) 컬러 공간으로 변환한 후 적색과 녹색의 색 차원을 유사한 색조로 사상(projection)하여 적록색맹이 보는 영상과 유사한 영상으로 변환하였다.

$$Q = R_Q P_R + G_Q P_G + B_Q P_B \quad (1)$$

where the vectors P_R , P_G , and P_B represent the three CRT primaries at maximum intensity. The coefficients R_Q , G_Q , and B_Q are weighting factors (ranging from 0 to 1) that determine the relative contribution of each primary to the generated color stimulus Q ^[22].

영상 변환은 MATLAB®을 이용하여 구현하였으며, Image Processing Toolbox를 사용하였다. Fig. 3의 결과는 컬러 영상을 적록색맹을 모사한 영상으로 자동 생성한 결과를 보여주는데 원본 컬러 영상에 비해 새롭게 생성된 적록색맹 영상의 배경과 위장군인이 상호 색조 대비가 커져 구분이 상대적으로 용이한 것을 확인할 수 있다.

5. 앙상블 딥러닝 학습 기반의 객체 인식

일반적인 앙상블 기법은 하나의 학습용 데이터 세트를 다수 딥러닝 모델로 학습시킨 후 산출된 다수 소결과 값을 배깅(Bagging), 부스팅(Boosting) 등의 방법으로 최종 결과 값을 예측하는 방식을 주로 사용한다. 즉 앙상블 딥러닝 방법은 약한 학습모델들을 잘 조합하여 강한학습 모델로 만드는 것을 목표로 한다. 즉 보통 성능의 모델들을 모아서 성능이 좋은 모델을 하나를 만드는 것이다.

배깅(Bagging)은 병렬적 방법으로 기본 데이터셋을 샘플링하여 n개의 데이터셋을 만들어서 n개의 모델을 학습시켜 최종 결정을 하는 모델이다. n개의 모델이 독립적으로 동시에 각각의 데이터셋을 학습할 수 있게 되어 속도가 빠르다. 부스팅(Boosting)은 직렬적 방법으로 첫 번째 모델이 기본 데이터셋을 그대로 학습하면 두 번째 모델은 전체 데이터를 학습하되 첫 번째 모델이 오류를 일으킨 데이터에 더 큰 중점을 두고 학습을 진행한다. 세 번째 모델은 앞의 두 모델이 맞추지 못한 데이터에 중점을 두고 학습을 진행한다. 앞 모델의 학습이 끝나야 뒷 모델이 그 결과를 기반으로 하여 가중치를 결정하고 학습을 할 수 있기 때문에 순차적으로 학습을 해야하기 때문에 상대적으로 속도가 느리다. 본 논문에서는 2개의 독립적인 학습용 데이터(컬러 원본 영상, 적록색맹 모사 영상)를 모두 적용하는 앙상블 딥러닝 기법을 고안하였다. 이렇게

하면 단일 학습용 데이터 세트의 앙상블 기법에 비하여 정확도가 높을 것으로 판단하였다. 앙상블 기법은 학습속도가 빠른 배깅(Bagging) 기법을 적용하였으며, 5개의 딥러닝 모델을 각 데이터 세트에 적용하여 학습을 진행하였다.

제안하는 앙상블 모델은 Fig. 4와 같이 2개의 학습용 영상 데이터(컬러 원본 영상, 적록색맹 모사 이미, Multi-input ensemble)를 독립적으로 입력받아 Table 2와 같은 5개의 모델로 나누어 학습하며 각각의 결과 값을 가중치(Weighted majority vote) 배깅 기법 통해 최종 인식결과를 예측하였다. 가중치는 적록색맹 영상 데이터 세트를 이용한 소결과 값에 더 큰 가중치를 부여하였다. 각 입력 데이터 세트는 5개의 딥러닝 모델을 통해 학습하도록 구성하여 총 10개의 소결과 값을 산출하고, 각 결과 값에는 가중치를 부여하여 최종 결과를 예측하도록 하였다.

학습 결과는 위장군인이 있는지 없는 지를 판별하는 Binary CrossEntropy 방식을 적용하였으며, 학습용 데이터는 위장군인이 포함된 영상과 없는 영상 1,419개, 검증용 데이터는 100장을 준비하였다. 딥러닝 모델은 3개 이상의 컨볼루션(Convolution) 레이어(3*3~7*7 Conv. + 3*3~5*5 Maxpooling or Averagepooling + Tanh)와 3개 이상의 Fully connected 레이어(ReLU or LeakyReLU or PReLU + Dropout)로 구성하여 학습을 수행하였다. 5개의 세부 모델의 구조는 Table 2와 같으며 다수 시행착오법을 통해 객체인식률이 높은 구조로 구성하였다. Learning rate는 0.01, Batch size는 20, Optimizer는 Adam을 적용하였다.

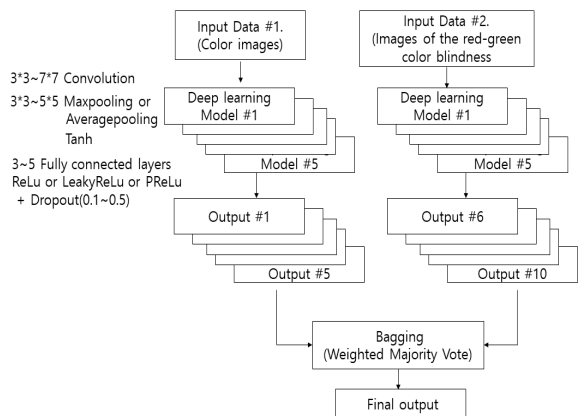


Fig. 4. Deep learning-based multi-input ensemble method^[1]

Table 2. Deep learning models(Model#1 ~5)

Model#1	Model#2	Model#3	Model#4	Model#5	
Input (3,240,360)	Input (3,240,360)	Input (3,240,360)	Input (3,240,360)	Input (3,240,360)	
Conv. Kernel(7,7)	Conv. Kernel(3,3)	Conv. Kernel(3,3)	Conv. Kernel(3,3)	Conv. Kernel(7,7)	
MaxPooling Shape(3,3)	AveragePooling Shape(3,3)	MaxPooling Shape(5,5)	MaxPooling Shape(3,3)	MaxPooling Shape(3,3)	
Tanh	Tanh	Tanh	Tanh	Tanh	
Conv. Kernel(5,5)	Conv. Kernel(5,5)	Conv. Kernel(5,5)	Conv. Kernel(3,3)	Conv. Kernel(5,5)	
MaxPooling Shape(3,3)	AveragePooling Shape(3,3)	MaxPooling Shape(5,5)	MaxPooling Shape(3,3)	MaxPooling Shape(3,3)	
Tanh	Tanh	Tanh	Tanh	Tanh	
Conv. Kernel(3,3)	Conv. Kernel(3,3)	Conv. Kernel(7,7)	Conv. Kernel(3,3)	Conv. Kernel(3,3)	×2
MaxPooling Shape(3,3)	AveragePooling Shape(3,3)				
Tanh	Tanh	Tanh	Tanh	Tanh	
Affine	Affine	Affine	Affine	Affine	
ReLU	ReLU	LeakyReLU Alpha(0.1)	LeakyReLU Alpha(0.1)	PRReLU	
Dropout (0.5)	Dropout (0.5)	Dropout (0.2)	Dropout (0.2)	Dropout (0.5)	
Affine	Affine	Affine	Affine	Affine	
ReLU	ReLU	LeakyReLU Alpha(0.1)	LeakyReLU Alpha(0.1)	PRReLU	
Dropout (0.5)	Dropout (0.5)	Dropout (0.2)	Dropout (0.2)	Dropout (0.5)	
Affine	Affine	Affine	Affine	Affine	
ReLU	ReLU	LeakyReLU Alpha(0.1)	LeakyReLU Alpha(0.1)	PRReLU	×2
Dropout (0.1)	Dropout (0.1)				
Affine	Affine	Affine	Affine	Affine	
Softmax					
Binary Cross Entropy					

제안된 방법의 성능비교를 위해 Table 3의 결과 중 방법1은 Inception-v3에 학습데이터로 컬러 영상만을

학습 후 성능을 평가하였고, 방법2는 적록색맹 모사 영상만을 학습데이터로 사용하여 Table 2와 같이 5개

Table 3. Classification accuracy of camouflaged soldier into similar background image

Methods		Data set	Accuracy	Avg.Precision	Avg.Recall	F-Measures
1	Inception-v3	Color images	56.00	54.06	56.00	56.60
2	Inception-v3	R-G color blindness images	72.00	72.14	72.00	71.95
3	Single-input Ensemble Method	Color images	65.00	74.36	65.00	61.27
4	Single-input Ensemble Method	R-G color blindness images	73.00	76.00	73.00	72.19
5	Multi-input Ensemble Method	Color images	67.00	67.86	67.00	66.59
6	Multi-input Ensemble Method	R-G color blindness images	76.00	76.37	76.00	75.91
7	Multi-input Ensemble Method	Color images + R-G color blindness images	86.00	86.53	86.00	85.95

딥러닝 모델로 학습하였다. 또한 방법3은 5개 딥러닝 모델에 컬러 영상을 학습 데이터로 방법4는 적록색맹 모사 영상만을 학습데이터로 사용하여 학습하였다. 방법5, 6, 7은 본 논문에서 제안하고자 한 모델을 적용하여 방법5는 컬러 영상만을 방법6은 적록색맹 영상을 사용하였고, 방법7은 컬러 영상과 적록색맹 영상을 모두 학습용 데이터로 사용하였다.

본 논문에서 제안한 방법7의 결과가 86 %로 가장 좋은 성능을 보였으며, 방법1과 2, 방법 3와 4, 방법 5와 6의 결과에서와 같이 같은 딥러닝 모델을 사용하였더라도 적록색맹 모사 영상을 학습데이터로 사용한 결과가 약 12~16 % 더 좋은 성능을 보였다. 이는 학습용 데이터의 질적 향상이 인식성능 향상과 상관관계가 있음을 보여주는 결과라 할 수 있다. 또한 같은 데이터 세트를 학습에 적용하였더라도 앙상블 학습 방법이 단일 모델 보다 약 5~10 %의 성능 향상을 보였다. Fig. 5는 인식결과의 성공 케이스와 실패 케이스의 예를 보여주는데 영상에서 위장군인 객체의 크기가 작아 일부 인식에 실패한 경우가 있는 것으로 판단되나 인식성공과 실패에 뚜렷한 상관관계는 식별되지는 않았다. 본 논문에서 제안한 방법이 전체적인 인식을 향상에는 기여하였으나, 여전히 적은 데이터로 인해 높은 수준의 인식을 달성은 어려웠다고 생각된다. 향후에는 데이터 증강(Data augmentation) 등의 방법을 통해 더 많은 수의 데이터를 확보하여 학습한다면 더욱 향상된 인식성능을 보일 것으로 예상된다.

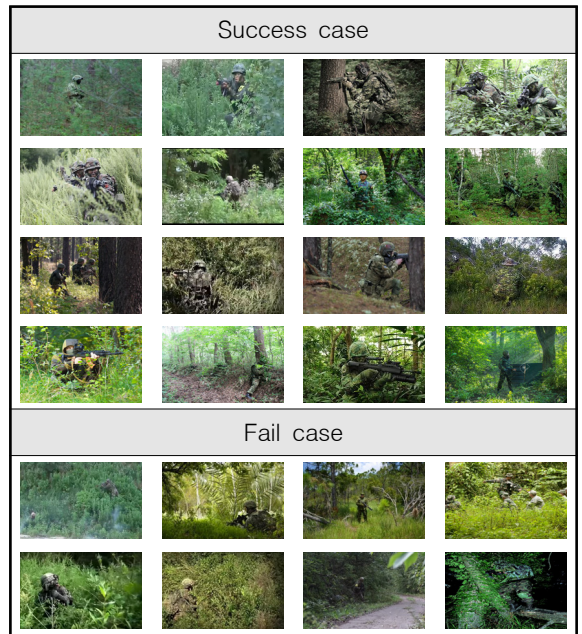


Fig. 5. Example images of the test result

6. 결론

본 논문에서는 유사배경 영상내에서 식별이 어려운 위장군인의 객체인식 성능 향상을 위한 적록색맹 모사 학습용 영상 생성기법과 다입력 앙상블 딥러닝 기

법을 제안하였다. 색맹인 사람에게는 일반인에게 유사하게 보이는 색들이 확연하게 다르게 보인다는 원리를 이용해서 적록색맹 영상 변환 알고리즘을 코드화하여 딥러닝 학습을 위한 모사 영상 데이터를 생성하였다. 생성된 적록색맹 모사 영상 데이터와 원본 컬러 데이터를 다입력으로 하는 앙상블 기법의 딥러닝 모델을 제안하여 위장군인의 인식률이 향상됨을 보였다. 향후 적록색맹 영상과 같이 학습용 데이터의 질을 더욱 높일 수 있는 데이터 생성기법을 발굴하여 다수 학습용 입력 데이터를 가지는 다입력 앙상블 기법을 발전시킬 필요가 있다고 생각된다.

References

[1] Choi K. H., "A Deep Learning-based Muti-input Ensemble Method for Performance Enhancement of Camouflaged Soldier Detection into Similar Background Image," 2019 KIMST Annual Conference Proceedings, pp. 1372-1373, 2019.

[2] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NIPS, 2012.

[3] M. D. Zeiler, R. Fergus, "Visualizing and Understanding Convolutional Networks," CoRR, 2013.

[4] K. Simonyan, A. Vedaldi, A. Zisserman, "Deep Fisher Networks for Large-scale Image Classification," NIPS, 2013.

[5] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," CVPR, 2016.

[6] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, A. Rabinovich, "Going Deeper with Convolutions," CVPR, 2015.

[7] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger, "Densely Connected Convolutional Networks," CVPR, 2016.

[8] Zhi-Hua Zhou, Ji Feng., "Deep forest: Towards an Alternative to Deep Neural Networks," arXiv preprint arXiv:1702.08835, 2017.

[9] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NIPS, 2012.

[10] R. Girshick, J. Donahue, T. Darrel, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," CVPR, 2014.

[11] R. Girshick., "Fast R-CNN," ICCV, 2015.

[12] Hanwang Zhang, Zawlin Kyaw, Jinyang Yu, Shih-Fu Chang, "PPR-FCN: Weakly Supervised Visual Relation Detection via Parallel Pairwise R-FCN," ICCV, pp. 4233-4241, 2017.

[13] Jifeng Dai, Yi Li, Kaiming He, Jian Sun, "R-FCN: Object Detection via Region-based Fully Convolutional Networks," CVPR, 2016.

[14] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," CVPR, 2016.

[15] Wei Liu, et. al., "SSD: Single Shot MultiBox Detector," ECCV, pp. 21-37, 2016.

[16] Hunmin Yang Ki-Jung Ryu Se-Yoon Oh, "Improving Deep Learning based Object Detection with Synthetic Data," KIMST Annual Conference Proceedings, pp. 1157-1158, 2019.

[17] Hunmin Yang Ki-Jung Ryu Se-Yoon Oh, "Visible-to-Infrared Image Translation using Adversarial Training," KIMST Annual Conference Proceedings, pp. 1436~1437, 2018.

[18] Soyeon Jin, "Detection of Weapon Systems using Deep Learning Object Detection," KIMST Annual Conference Proceedings, 2019.

[19] Se-Yoon Oh Ki-Jung Ryu, Hunmin Yang, "Experimental Design with Synthetic and Real Datafor Efficient Object Detection," KIMST Annual Conference Proceedings, pp. 1153-1154, 2019.

[20] C. Szegedy, et. al., "Rethinking the Inception Architecture for Computer Vision," CVPR, 2016.

[21] JM Bosten, JD Robinson, G Jordan, JD Mollon, "Multidimensional Scaling Reveals a Color Dimension Unique to 'Color-Deficient' Observers," Current Biology, 2005.

[22] Hans Brettel, et. al., "Computerized Simulation of Color Appearance for Dichromats," J. Opt. Soc. Am., Vol. 14, No. 10, 1997.