

자동차 번호판 위치 인식을 위한 적대적 도메인 적응 모델 경량화

조 항 재*, 김 정 현*, 전 경 구^o

Lightweight Adversarial Domain Adaptive Model for License Plate Detection

Hangjae Cho*, Jeonghyeon Kim*, Kyungkoo Jun^o

요 약

딥러닝을 활용한 자동차 번호판 위치 인식 기술은 자동차 번호판 인식 문제에서 매우 중요한 선행 작업이다. 그러나 딥러닝 모델의 특성상, 학습된 환경과 다른 환경에서 추론할 때 성능이 저하되는 문제가 존재한다. 이를 해결하기 위해 적대적 도메인 적응 기술을 활용하여 성능을 향상시킬 수 있다. 하지만, 적대적 도메인 적응 기법을 RetinaNet에 적용한 DA(Domain Adaptation) RetinaNet 모델을 사용하여 실험한 결과, 사용된 자동차 번호판 데이터의 크기가 모델 크기에 비해 작아서 성능 향상이 나타나지 않았다. 이러한 문제를 극복하기 위해 본 연구에서는 데이터 크기에 맞게 DA RetinaNet을 경량화하고 사전 학습을 활용하여 자동차 번호판 위치 인식 문제의 성능을 향상시키는 방법을 제안한다. 제안된 방법을 적용하면, 소스 도메인 데이터만을 사용하여 학습 후 타 도메인에서의 추론을 수행했을 때와 비교해 F1 Score 기준으로 최대 45% 향상 및 mAP@0.5 기준으로 최대 91% 향상을 달성할 수 있었다.

Key Words : Deep learning, license plate detection, adversarial domain adaptation

ABSTRACT

The technology of car license plate location recognition using deep learning is a crucial prerequisite in solving the car license plate recognition problem. However, due to the inherent characteristics of deep learning models, there exists a challenge of performance degradation when making inferences in environments different from the ones they were trained on. To address this issue, performance can be enhanced by leveraging adversarial domain adaptation techniques. However, the experiments conducted using the DA (Domain Adaptation) RetinaNet model, which applied adversarial domain adaptation techniques to RetinaNet, revealed no performance improvement. This lack of improvement was attributed to the fact that the size of the car license plate data used was smaller compared to the model size. To overcome these challenges, this study proposes a method to enhance the performance of the car license plate location recognition problem by customizing DA RetinaNet to match the dataset size and utilizing pre-training. Applying the proposed method resulted in up to

※ 이 논문은 인천대학교 2023년도 자체연구비 지원(2023-0037)에 의하여 연구되었습니다

• First Author : Dept. of Embedded Systems Engineering Incheon National University, hangjae99@inu.ac.kr, 정회원

o Corresponding Author : Dept. of Embedded Systems Engineering Incheon National University & Energy Excellence & Smart City Lab., kjun@inu.ac.kr, 정회원

* Dept. of Embedded Systems Engineering Incheon National University, holssy2@gmail.com

논문번호 : 202308-052-C-RN, Received August 18, 2023; Revised October 6, 2023; Accepted October 10, 2023

a 45% improvement based on the F1 Score and up to a 91% improvement based on mAP@0.5 compared to the performance achieved when only source domain data was used for training, and inference was conducted in other domains.

1. 서 론

딥러닝으로 번호판 위치를 정확하게 파악하는 것은 번호판 번호 인식 과정 중 선행되는 작업이다.^[1] 자동차 번호판 인식은 주변 환경과 다양한 숫자와 문자가 더불어 등장하는 복잡한 영상에서 작동해야 하기 때문이다. 또한, 번호판이 차지하는 공간이 제한적이기 때문에 입력된 사진을 가공하지 않고 번호판 내의 문자나 숫자를 인식하기에 어려울 수 있다. 이 문제를 해결하기 위해 번호판 인식 이전에 자동차 번호판의 위치를 특징하여 해당 영역에서 번호판을 인식하는 방법을 적용한다. 이로써 영상에서 번호판이 아닌 다른 영역의 숫자나 문자를 인식하는 문제를 방지하고, 온전한 번호판 내의 숫자와 문자를 인식할 수 있게 된다.

하지만, 딥러닝을 사용하여 번호판 위치 인식을 진행할 때, 학습에 사용한 영상과 다른 환경의 영상을 이용할 경우 성능이 떨어지는 문제가 발생한다. 이 문제는 딥러닝 모델이 학습한 환경에 특화되기 때문에 발생한다. 예를 들어, 한국의 번호판의 위치를 인식할 수 있는 모델을 사용하여 미국 등 외국의 영상을 입력으로 주었을 때, 이를 탐지하는 성능이 매우 떨어진다. 이를 해결하기 위해 기존의 번호판 위치 인식 모델은 새로운 환경에 대해서 매번 학습을 수행하여야 했다. 그러나, 새로운 환경이 나타날 때마다 학습 데이터를 만드는 작업은 많은 인력과 시간을 필요로 하는 문제가 있다.

딥러닝에서는 학습에 사용되지 않는 다른 환경의 영상에서도 추가적인 노력 없이 기존 모델을 사용하여 높은 성능을 얻는 방법으로 도메인 적응(Domain Adaptation) 방법^[2,3]이 제안되었다. 여기서 도메인(Domain)이란 데이터의 특성과 분포를 결정하는 요소들의 집합을 말한다. 즉, 도메인 적응은 서로 다른 두 도메인 간의 데이터 분포를 유사하게 만들어서 정확도를 높이는 방법이다. 자동차 번호판 위치 인식 문제에서 도메인 적응을 적용하면, 새로운 환경마다 데이터를 생성하는 번거로움을 줄이고 학습 시간을 단축할 수 있다.

자동차 번호판 위치 인식 문제에 준지도 도메인 적응(Semi-Supervised Domain Adaptation) 방식을 적용

한 객체 탐지 모델인 DA Retinanet^[4]을 경량화하여 추가 데이터 가공 없이 성능 향상을 이룰 수 있었다. 데이터는 그림 1과 같은 한국, 미국, 필리핀 3가지의 도메인에서 각각 1,000장씩 사용하여 진행하였다. DA Retinanet 모델을 그대로 적용하였을 때, 성능 향상이 나타나지 않았으며 오히려 손실(Loss)이 높아져 정상적인 학습이 이루어지지 않았다. 이를 해결하기 위해 DA Retinanet의 Loss 함수를 변경하고 모델의 백본(Backbone)과 도메인 헤드(Domain Head)를 경량화하였다. 또한, 사전학습을 진행하여 성능을 더 끌어올렸다. 경량화된 모델로 학습을 진행한 결과, 소스 데이터만으로 추론을 진행하였을 때보다 F1 Score 기준 최대 45% 성능 향상과, mAP@0.5 기준 최대 91% 성능 향상을 달성할 수 있었다.

본 논문의 주요 내용은 다음과 같다. 첫째, 기존의 적대적 도메인 적응 모델을 자동차 번호판 위치 인식 데이터에 맞게 경량화하였다. 둘째, 실험을 통해 사전 학습이 적대적 도메인 적응 학습의 성능 향상에 영향을 준다는 것을 확인하였다. 셋째, 실험을 통해 제안한 경량화 모델과 사전 학습이 자동차 번호판 위치 인식 성능을 F1 Score 기준 최대 45%, mAP@0.5 기준 최대 91% 향상한다는 것을 확인하였다.

본 논문은 다음과 같이 구성된다. 2 장에서는 도메

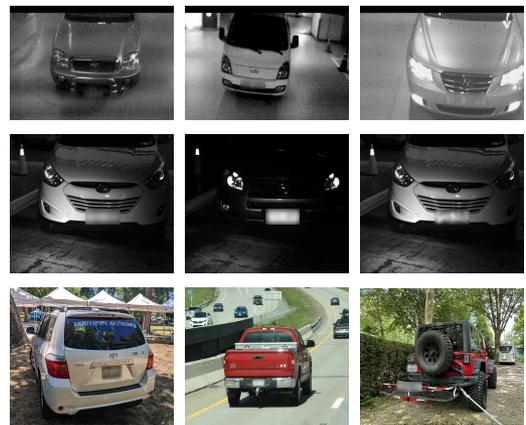


그림 1. 연구에서 사용된 3개 국가의 자동차 번호판 원본 사진 예시. 위에서부터 한국, 필리핀, 미국
Fig. 1. Examples of original photos of license plates from the three countries used in the study. From top to bottom: Korea, Philippines, USA

인 학습에 대한 간략한 설명과 기준에 제안된 적대적 도메인 학습 모델에 대한 설명을 진행한다. 3 장에서는 기존 모델이 자동차 번호판 문제에서 어떤 문제를 가지고 있는지 파악하고, 이를 해결하기 위해 진행한 경량화 방법과 성능을 더 향상하기 위해 실험적으로 시도한 방법들을 소개한다. 4 장에서는 3 장에서 소개된 모델과 실험적으로 시도한 방법들에 대한 실험 결과를 설명한다. 마지막으로 5장에서는 결론으로 마무리 짓는다.

II. 관련 연구

본 장에서는 도메인 적응 기법에 대한 간략한 소개와 기존 논문의 DA RetinaNet에 대해 설명하고자 한다. 1절에서는 도메인 적응 기법의 개요와, 본 논문에서 사용한 적대적 도메인 적응 기법과 대표적인 예인 DANN(Domain Adaptation Neural Network)^[5]에 대해 간략히 설명한다. 2절에서는 본 논문에서 사용한 DA RetinaNet 모델의 구조와 진행 과정을 간략하게 설명한다.

2.1 도메인 적응 기법

도메인 적응은 기계 학습에서 한 도메인의 데이터로 학습한 모델을 다른 도메인으로 일반화하는 기술을 말한다. 도메인은 데이터의 분포와 특성을 나타내며, 도메인 간의 차이로 인해 특정 도메인에서 학습된 모델이 다른 도메인에서 성능이 저하되는 문제가 발생할 수 있다. 도메인 적응은 이러한 도메인 간의 차이를 줄이거나 보완하여 모델의 일반화 능력을 향상시켜 타겟 도메인(Target Domain)에서의 성능을 높여 준다.

도메인 적응 기법은 타겟 레이블의 유무와 학습 방법에 따라 분류할 수 있다. 우선 도메인 적응 기법은 레이블에 유무에 따라 3가지로 분류할 수 있다.^[6] 이는 비지도 도메인 적응 방법(Unsupervised Domain Adaptation)^[7], 지도 도메인 적응 방법(Supervised Domain Adaptation)^[8], 준지도 도메인 적응 방법(Semi-Supervised Domain Adaptation)^[9]이 해당된다. 또한, 학습 방법에 따라 3가지로 분류될 수 있다. 불일치 기반 방법(Discrepancy-Based Method)^[10], 적대적 기반 방법(Adversarial-based Method)^[11], 자기 지도 학습 기반 방법(Self-Supervision-based Method)^[12]로 나눌 수 있다.

본 연구에서 사용할 방법은 적대적 기반 방식의 준지도 도메인 적응 방법이다. 준지도 도메인 적응 방식

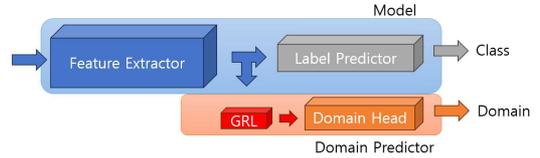


그림 2. 분류 문제에서의 적대적 도메인 적응 기법인 Domain Adaptation Neural Network(DANN)의 구조
Fig. 2. Structure of Domain Adaptation Neural Network (DANN), an adversarial domain adaptation in classification

은 정답 레이블이 있는 소스 도메인(Source Domain)과 정답 레이블이 없는 타겟 도메인으로 구성된 데이터로 학습하는 방식이다. 적대적 기반 방식은 GAN(Generative Adversarial Network)^[13]의 알고리즘을 이용하여 두 도메인을 한 도메인으로 일반화하는 방식이다. 적대적 기반 방식은 해당 알고리즘을 이용하여 특징을 추출하는 특징 추출기(Feature Extractor)와 입력이 어느 도메인에 속하는지 구분해주는 도메인 헤드로 구성된다. 이때, 도메인 헤드는 입력된 데이터가 어느 도메인인지 구분할 때 나온 손실에 음의 상수를 곱하여 특징 추출기로 넘겨 도메인 판별을 방해함으로써 두 도메인을 한 도메인으로 일반화 시키는데 도움을 준다.

DANN(Domain Adversarial Neural Network)은 그림 2와 유사한 형태로 구성되며, 도메인 적응 분야에서 최초로 등장한 기법이다. DANN은 특징 추출기와 도메인 헤드를 연결하는 과정에서 기울기 반전 레이어(Gradient Reversal Layer)^[7]를 활용하여 적대적 도메인 적응을 수행한다. 이러한 기울기 반전 레이어는 순전과 과정에서는 입력을 정상적으로 도메인 헤드로 전달하지만, 역전과 과정에서는 기울기에 음의 상수를 곱하여 학습을 방해한다. 즉, 두 도메인을 구별하는 학습을 방해하여 두 도메인 간의 특징을 구별할 수 없게 만들어, 두 도메인을 한 도메인으로 여기게 하여 도메인 적응을 효과적으로 수행하게 한다.

2.2 DA RetinaNet 모델

DA RetinaNet을 설명하기 전, 사용한 모델인 RetinaNet^[14]에 대해 설명한다. RetinaNet은 FPN(Feature Pyramid Networks)^[15]의 구조를 사용하고, 가중치 공유라는 기법을 사용하는 것이 특징이다. 입력된 이미지는 특징 추출기이자 백본인 ResNet^[16]에 의해 특징이 추출된다. 이렇게 추출된 특징들을 FPN으로 넘겨주는데, FPN 구조는 이미지의 다양한 크기의 특징을 사용하여 물체를 탐지하는 네트워크 구조이며 이는 높은 정확도를 얻을 수 있도록 해준다.

FPN에는 객체의 위치를 예측하는 bounding box head와 classification head가 subnet으로 결합되어 있어 객체의 위치 인식과 분류가 가능하다. FPN에 사용된 가중치 공유 방식은 동일한 특성 맵에 가중치를 공유하게 함으로써 학습 시간을 감소시켜 모델의 추론 시간과 크기를 감소시켜준다.

DA RetinaNet은 적대적 학습 방식인 DANN을 RetinaNet 모델에 적용하여 도메인 적응 학습을 수행한 모델이다. DANN과 동일하게 RetinaNet의 백본인 ResNet에 도메인 헤드를 붙임으로서 RetinaNet에 적대적 도메인 적응 학습을 진행한다. 다만 기존의 DANN 방법과 다르게 도메인 헤드를 3개로 늘려 백본인 ResNet의 마지막 3개의 Residual Block에서 나온 결과를 각각 도메인 학습에 사용한다. 모델의 입력으로 레이블이 있는 소스 도메인과 레이블이 없는 타겟 도메인이 동시에 입력되어 진행된다. 만일 입력된 사진이 소스 도메인일 경우, 백본에서 나온 결과가 도메인 헤드와 FPN에 동시에 전달되어 모델의 손실과 도메인 손실(Domain Loss)를 얻는다. 반면, 입력된 사진이 타겟 도메인일 경우, 백본의 통해 나온 결과는 FPN에 전달되지 않고 도메인 헤드로만 전달되어 도메인 손실만 구하게 된다. 이렇게 구한 모델의 손실과 두 도메인 손실을 역전파로 전달하면서 학습을 진행한다.

III. 제안 방법

본 연구에서 자동차 번호판 데이터를 이용해 DA RetinaNet을 학습한 결과, 성능 하락 문제가 나타났다. 이러한 문제는 학습 및 추론에 사용된 데이터의 차이 때문으로 예상된다. DA RetinaNet은 크기가 큰 데이터 입력을 고려하여 설계되었기 때문에, 크기가 작은 데이터를 사용할 때는 성능 향상이 어려운 것으로 보인다. 실제로, DA RetinaNet에서 사용된 데이터는 박물관 3D 모델링 연구¹⁷⁾에서 파생된 데이터로, 소스 데이터는 실제 박물관 사진을, 타겟 데이터는 3D 모델링을 통해 생성된 데이터를 활용했다. 이 데이터셋은 가로 1280 픽셀, 세로 720 픽셀의 크기를 가지며, 2190장의 소스 데이터와 100,769장의 타겟 데이터로 구성되었다. 또한, 클래스(Class) 개수는 16개로, 본 연구에서 다루는 데이터와는 다르게 다양한 물체의 이미지를 포함하고 있다. 따라서 이 모델을 그대로 사용하는 것은 자동차 번호판 위치 추론 문제에는 부적합하다고 판단된다.

문제를 해결하기 위해, DA RetinaNet 백본을 경량

화하는 방안을 제안한다. DA RetinaNet에서 사용된 특징 추출기인 ResNet101은 매우 크고 복잡한 모델로, 작은 데이터셋에서 학습할 때 오히려 원활한 학습이 이루어지지 않을 수 있다. 더구나, 본 연구에서 활용되는 이미지는 384 x 384 크기로 작고, 제로 패딩(zero padding)이 적용되어 있으며, 클래스 개수도 자동차 번호판 1개로 적다. 따라서 이러한 조건을 고려하여, DA RetinaNet에서 사용된 ResNet 모델을 가장 경량화된 형태인 ResNet18로 변경하였다. ResNet18은 ResNet 모델 중에서 작은 데이터셋에서 효과적으로 학습할 수 있도록 설계된 모델이다.

모델을 추가적으로 경량화 시키기 위해 기존 백본의 도메인 헤드의 개수를 줄이는 방안을 제안한다. 특징 추출기인 ResNet이 경량화되어, 다수의 도메인 헤드를 이용해서 진행하면 기존보다 적어진 계층들에게 도메인 손실이 증첩되어 정상적으로 학습되지 않는다. 본 연구에서는 기존에 3개였던 도메인 헤드를 2개로 줄여 사용하더라도 도메인 손실이 발산하는 문제가 나타나 도메인 헤드 1개만을 사용하였다.

도메인 헤드의 개수를 1개로 줄였을 때, 해당 도메인 헤드를 ResNet의 블록 3에 붙여 진행하는 것을 제안한다. 도메인 헤드를 마지막 블록에 붙이게 되면 백본인 ResNet의 계층이 연장되는 효과가 나 모델이 커지기 때문에 경량화한 장점이 사라진다. 또한, 블록 4, 5에 붙일 경우, 기존의 DA Retinanet에 사용한 데이터보다 작은 데이터를 사용하였기 때문에 해당 계층에서 나온 특징의 크기가 작아 학습하는데 어려움이 발생하였다. 이러한 이유로, 특징 추출기인 ResNet의 블록 3에 붙여서 진행하였다. 실제로 블록 5에만 붙여서 학습을 진행하였을 때 도메인 손실이 발산하면서 수렴이 되지 않아 학습이 정상적으로 이루어지지 않았다.

제안한 모델은 그림 3와 같으며 백본의 블록 3에 도메인 헤드를 기울기 반전 레이어를 통해 연결하였다. 앞서 설명한 대로 모델을 경량화 하기 위해 DA RetinaNet보다 작은 백본을 사용하였기 때문에 블록 3에만 도메인 헤드를 붙였다. 도메인 헤드 경량화 또한 진행하였다. 기존과 동일하게 2개의 합성곱 계층으로 구성되지만 줄어든 이미지 크기에 맞추어 입력과 출력의 채널 수를 조절하였다. 도메인 헤드의 추론하는 부분도 경량화를 진행하여 Adaptive Average Pooling 후 합성곱을 통해 값을 얻었던 기존과 달리, 바로 Flatten을 진행한 후 Linear를 통해 결과를 구한다.

도메인 손실을 구하기 위한 손실함수는 Cross Entropy Loss¹⁸⁾를 사용한다. 기존 DA RetinaNet은

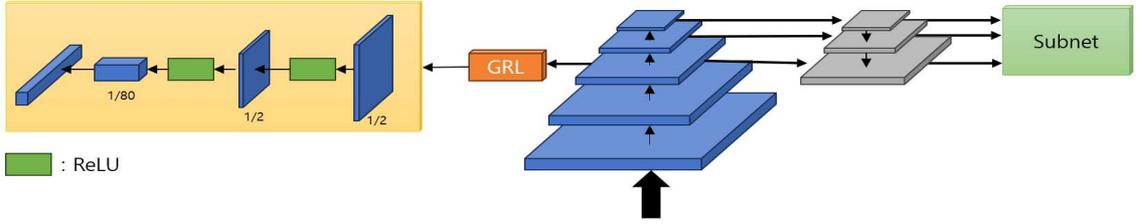


그림 3. 연구에서 제안한 DA RetinaNet을 경량화한 모델
 Fig. 3. A lightweight model of the DA RetinaNet proposed in the study

Retinanet 논문에서 제안되었던 Sigmoid Focal Loss[14]을 사용하였다. Sigmoid Focal Loss는 Cross Entropy Loss를 개량한 손실함수로, 클래스(Class) 개수의 불균형성을 보완하기 위해 나온 함수이다. 하지만 자동차 번호판 위치 인식에 사용한 데이터는 클래스가 번호판 하나이기 때문에 Sigmoid Focal Loss를 사용하지 않아도 된다. 사용한 Cross Entropy Loss는 식 (1)과 같다.

$$Loss = -\frac{1}{n} \sum_{i=1}^n \sum_{c=1}^C L_{ic} \log p_{ic} \quad (1)$$

식에서 n 은 데이터의 개수를 의미하고, C 는 class의 개수, L 은 실제 값, P 는 실제 확률을 나타낸다.

또한, 성능 향상을 위해 소스 도메인으로 사전 학습을 진행하였다. DA RetinaNet일 경우, 사전 학습은 Model Zoo에 있는 사전 학습 모델을 사용하였다. 하지만 해당 모델은 학습하는 사진과 관련 없는 데이터로 학습되어, 학습 대상의 데이터로 추가 사전 학습을 진행하였을 때보다 추론의 성능이 떨어진다. 따라서 해당 사전 학습 모델과 더불어 소스 데이터로 사전 학습을 선행하여 진행하였다. 실제로 소스 데이터로 학습이 된 사전 학습 모델에 적대적 도메인 적응 방식을 추가하니 사전 학습 없을 때보다 성능 향상이 이루어졌다.

적대적 도메인 적응 방법의 성능을 향상시키기 위해서는 도메인 헤드의 정확도를 향상시켜야 한다. 적대적 도메인 적응 방법은 GAN에서 영감을 받아 개발되었기 때문에, 도메인 헤드가 정확하게 도메인을 분류할수록 기울기 반전 레이어를 통해 두 도메인 간의 구분을 어렵게 만들 수 있다. 즉, 도메인 헤드가 정확하게 도메인을 분류해줄수록 두 도메인간의 구분이 어려워져서 타겟 도메인에서의 추론 성능을 향상시킬 수 있다.

도메인 헤드의 정확도를 높이기 위해 Residual

block을 넣었으나 성능 향상이 크게 나타나지 않았다. Residual block은 그림 3의 도메인 헤드의 두 합성곱 레이어에 적용하였다. Residual block을 적용하여도 성능 향상이 두드러지게 나타나지 않는 이유는, Residual이 적용된 block의 크기가 충분히 크지 않고 블록 내부의 합성곱 크기도 작아 실질적으로 큰 성능 향상이 이뤄지지 않았기 때문으로 보인다.

IV. 실험

본 장의 1절에서는 실험에 사용한 데이터와 하이퍼 파라미터의 세팅과 실험 결과를 설명한다. 데이터에서는 도메인별 장 수와 데이터 전처리 과정을 설명한다. 또한 하이퍼 파라미터 문단에서는 실제 학습에 적용된 하이퍼 파라미터와 평가 지표를 설명한다. 이후 3장에서 제안한 방법을 적용한 모델로 진행한 결과를 설명한다.

4.1 데이터 셋 및 하이퍼 파라미터

다른 도메인간의 적대적 도메인 적응 학습을 위해 총 3개의 도메인 데이터를 수집하였다. 수집된 도메인은 한국, 필리핀, 미국으로, 각 도메인별 학습에 사용된 장 수는 다음 표 1과 같다. 수집된 한국 데이터는 건물의 지하 주차장 출입구에 설치된 폐쇄회로 카메라(CCTV)에서 찍힌 사진이며, 사진에는 많은 노이즈가 있다. 필리핀 데이터는 어두운 환경으로, 동일한 위치의 외부 주차장에서 찍은 사진들이다. 이는 한국 데이터와는 달리 노이즈가 없는 깨끗한 사진이며, 내

표 1. 연구에서 사용한 도메인 별 train과 test의 이미지 장수
 Table 1. The number of images of train and test for each domain used in the study

Data	Korea	Philippines	U.S.A
Train	800	800	800
Test	200	200	200

부가 아닌 외부에서, 어두운 환경에서 진행하였기 때문에 환경적인 차이가 커 선정하였다. 미국의 경우, 한국과 필리핀의 차량에 있는 번호판의 위치와는 다른 곳에 번호판이 위치된 사진도 존재하였으며, 타 국가의 번호판보다 여러 그림 및 다양한 문양이 번호판에 있어 선정하였다.

수집된 데이터들은 학습에 동일한 입력으로 사용하기 위해 학습 전 전처리를 진행하였다. 우선, 안정적인 인식을 위해 컬러 이미지는 흑백 이미지로 전환하였다. 흑백 이미지로 변경하는 이유는, 조명에 의한 영향을 최대한 적게 받기 위함이다. 이후, 각 사진은 384 x 384 픽셀 크기로 비율에 맞추어 크기를 변경하고, 남은 공간은 제로 패딩(zero padding)으로 진행하였다.

학습에 이용한 초기 파라미터와 평가 기준은 다음과 같다. 학습 시 배치(batch)의 크기는 소스 데이터와 타겟 데이터 모두 12, 최적화 알고리즘은 Adam^[19]을 사용하였다. 또한, 초기 학습률(LR, Learning Rate)은 0.00001, 스케줄러(Scheduler)는 StepLR을 사용하였다. 데이터 증강을 위해 좌우 랜덤 플립(Left-Right Random Flip)과 이미지 Resize를 적용한다.

비교 기준은 객체 탐지 평가 지표 중 F1 Score과 mAP@0.5를 통해 평가를 진행하였다. F1 Score의 경우, 정밀도(Precision)과 재현율(Recall)을 동시에 고려하기 때문에 단순히 정확도(Accuracy)만으로 판별하는 것보다 정확하게 모델의 성능을 파악할 수 있다. 또한 mAP@0.5를 통해 모델이 추론한 결과들이 실제 정답의 위치를 얼마나 정확하게 위치하는지 파악할 수 있어 해당 평가 지표를 사용하였다.

4.2 실험 결과

소스 데이터로만 학습을 진행하고 타겟 도메인을 추론하였을 때에는 F1 score와 mAP@0.5가 각각 최대 0.91, 0.89 최소 0.20, 0.11점의 결과가 나타났다. 학습은 소스 데이터로 150 에포크(epoch) 진행하였으며, 그때의 타겟 도메인의 추론 결과는 다음 표 2와 같다. 본 연구에 이루어진 모든 실험의 결과는 소스 데이터를 추론하였을 경우 F1 Score와 mAP@0.5는 모두 0.96점 이상 유지하였기에 표에서 제외하였다.

사전 학습 없이 소스 데이터와 타겟 데이터를 가지고 개량된 모델로 실험하였을 경우 F1 Score 기준 최대 16%, mAP@0.5는 최대 44% 향상되었다. 실험은 150 에포크를 진행하였으며, 결과는 표 3과 같다. 다만, 소스 도메인이 한국이었을 때 거의 변화가 없는 모습이 나타났다. 이는 한국의 데이터가 다른 두 도메

인 데이터를 추론할 때 이미 높은 점수를 받았기 때문에 큰 향상이 이루어지지 않은 것으로 보인다. 반면 소스 데이터가 미국과 필리핀일 경우, F1 Score는 최대 16%, mAP@0.5는 최대 44% 향상된 모습이 나타났다. 하지만, 해당 상승분은 기준에 낮은 점수를 기록한 점수의 향상이기 때문에 실제 점수로는 0.6에서 0.1사이로 큰 향상이 이루어지지 않았음을 알 수 있다.

사전학습 관련해서 학습을 진행하였을 때, 모든 결과에서 F1 Score가 상승하였으며 F1 Score는 최대 45%, mAP@0.5는 91%까지 향상되었다. 사전 학습을 진행하기 위해 100 에포크까지 소스 데이터만으로 학습을 진행하고, 이후 50 에포크 동안 적대적 도메인 학습을 진행하였으며 결과는 표 4와 같다. 전체적으로 사전학습을 진행하지 않았을 때보다 모든 결과가 향상되었으며, 대부분의 경우 처음부터 진행하였던 표 3

표 2. 소스 도메인만 가지고 학습하였을 때의 타겟 도메인 추론 결과
Table 2. Target domain inference result when learning with only the source domain

Source Domain	Target Domain	F1 Score	mAP@.5
Korea	Philippines	0.7033	0.5717
	U.S.A	0.9119	0.8994
Philippines	Korea	0.2965	0.1356
	U.S.A	0.6555	0.4829
U.S.A	Korea	0.2037	0.1139
	Philippines	0.5465	0.4057

표 3. 제안한 경량화 모델로 사전 학습 없이 150 에포크 진행하였을 때의 결과(괄호 안은 표 2의 결과와의 차이)
Table 3. Results when 150 epochs were run without prior training with the proposed lightweight model (differences from the results in Table 2 in parentheses)

Source Domain	Target Domain	F1 Score	mAP@.5
Korea	Philippines	0.7289 (+3%)	0.6044 (+5%)
	U.S.A	0.9554 (+4%)	0.9542 (+6%)
Philippines	Korea	0.3445 (+16%)	0.1956 (+44%)
	U.S.A	0.787 (+20%)	0.6684 (+38%)
U.S.A	Korea	0.2304 (+13%)	0.1302 (+14%)
	Philippines	0.6055 (+10%)	0.4674 (+15%)

표 4. 제안한 경량화 모델로 사전 학습을 포함해 150 에포크 진행한 결과(괄호 안은 표 2의 결과와의 차이)
Table 4. Results of 150 epochs including pre-learning with the proposed lightweight model (differences from results in Table 2 in parentheses)

Source Domain	Target Domain	F1 Score	mAP@.5
Korea	Philippines	0.753 (+10%)	0.6248 (+9%)
	U.S.A	0.9443 (+3%)	0.9645 (+7%)
Philippines	Korea	0.4302 (+45%)	0.2597 (+91%)
	U.S.A	0.7729 (+17%)	0.6423 (+33%)
U.S.A	Korea	0.3217 (+57%)	0.1925 (+69%)
	Philippines	0.7065 (+29%)	0.5995 (+47%)

의 실험보다 높은 차이로 향상되었다. 가장 많이 향상된 점수 역시 사전 학습을 진행하지 않았을 때 가장 많이 향상된 정도보다 2배 이상 향상되는 결과가 나왔다. 이는 적대적 도메인 적응 학습을 하기 전 사전 학습으로 자동차 번호판의 모습에 대해 미리 학습을 진행한 후, 도메인 간 거리를 좁히니 성능이 더 올라갔음을 알 수 있다.

도메인 헤드의 성능을 높이기 위해 Residual block을 사용해 진행하였을 때에는 대부분의 경우 성능이 Residual block이 없는 실험보다 낮음을 알 수 있다. Residual block 훈련 또한 사전 학습을 진행하였지만, 도메인 헤드의 정확성을 높여 점수가 더 올라갈 것이라는 기대에는 못 미치는 결과를 얻었으며 결과는 표 5과 같다. 일부 상승폭이 큰 결과가 있었지만, 1% 정도의 미비한 차이이며, 해당 결과는 원래도 높은 정확도를 보이는 결과이기 때문에 전체적으로 학습이 덜 되는 것으로 파악된다. 이는 도메인 헤드의 경량화가 진행된 Residual이 적용된 block이 충분히 깊지도 않을뿐더러 도메인 헤드 자체의 크기도 작아 Residual이 효능이 있지 않았던 것으로 보인다. 따라서 자동차 번호판 위치 인식 적대적 도메인 적응 모델의 도메인 헤드에 Residual block을 진행하는 것은 부적절하다고 판단된다.

도메인 헤드의 위치를 변경하여 진행하였을 때 도메인 손실이 발산하여 정상적인 학습이 이루어지지 않았다. 예를 들어, Retinanet에서 FPN과 연결되는 백본 Resnet의 블록은 3, 4, 5가 존재한다. 이때 블록 5

에만 도메인 헤드를 붙여 진행하였을 경우, 그림 4과 같이 도메인 손실이 발산하였다. 도메인 손실이 발산하면서 오히려 소스 데이터의 학습에도 영향을 주어 자동차 번호판 위치 인식을 불가능하게 만들었다. 도메인 헤드는 특징 추출기에 추가로 연결되어 도메인 헤드의 손실이 발산하게 되면 특징 추출기의 손실도 같이 발산되기 때문에 기존 학습도 진행되지 못하게 된다. 이외에 블록 4에도 진행하였지만, 블록 5와 같이 도메인 손실이 발산하여 본 논문의 모델은 특징 추출기의 블록 3에 도메인 헤드를 붙여 진행하는 것이 적절하다고 보인다.

표 5. 사전 학습과 Residual을 적용하여 150 에포크 진행했을 때의 결과(괄호 안은 표 2의 결과와의 차이)
Table 5. Results when 150 epochs were performed by applying Residual, including prior learning (differences from results in Table 2 in parentheses)

Source Domain	Target Domain	F1 Score	mAP@.5
Korea	Philippines	0.719 (+2%)	0.5931 (+3%)
	U.S.A	0.9542 (+4%)	0.9782 (+8%)
Philippines	Korea	0.3333 (+12%)	0.196 (+44%)
	U.S.A	0.7572 (+15%)	0.6179 (+27%)
U.S.A	Korea	0.2046 (+0%)	0.1139 (+0%)
	Philippines	0.6319 (+15%)	0.4889 (+20%)

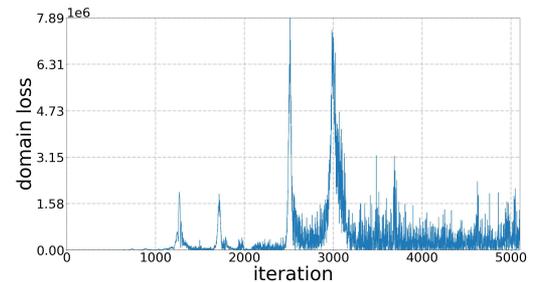


그림 4. 블록 5에 도메인 헤드를 붙였을 때 도메인 손실
Fig. 4. Domain loss when attaching domain head to block 5

V. 결론

본 논문에서는 자동차 번호판 위치 인식에 맞는 경

량화된 적대적 도메인 적응 방식을 제안하였다. DA RetinaNet의 적대적 도메인 적응 방식은 본 논문의 데이터 세트에 비해 모델이 커 도메인 손실이 수렴하지 못하는 문제가 발생하였다. 이를 해결하고자 도메인 헤드를 경량화하고 사전 학습을 이용하여 성능을 향상시켰으며 실험을 통해 경량화된 모델은 소스 도메인만으로 학습하였을 때보다 F1 Score 기준 최대 45%, mAP@0.5 기준 최대 91% 성능 향상을 이루어 낼 수 있음을 증명하였다.

추후 연구에서는 적대적 도메인 학습의 효율을 향상시키기 위해 도메인 헤드의 정확도를 높이거나 손실 함수를 조정하는 방안을 연구해보고자 한다. 적대적 기반 방식 특성상 도메인 헤드의 정확도가 높아질수록 두 도메인간 성능 향상이 기대되기 때문이다. 더불어 현재 사용 중인 Cross Entropy Loss는 주로 원 핫 인코딩에 적용되지만, 기울기 반전 레이어를 통한 학습 특성으로 인해 적대적 도메인 적응 문제에는 더 적합한 최적의 손실 함수가 존재할 것이라고 보인다.

References

[1] W. Kim, S. Cho, N. T. Phuong, N. D. Dong, H. K. Lee, and K. Lee, "Licence plate recognition system with image processing and deep learning," *J. KIIS*, vol. 31, no. 5, pp. 429-435, 2021. (<https://doi.org/10.5391/JKIIS.2021.31.5.429>)

[2] M. Wang and W. Deng, "Deep visual domain adaptation: A survey," *Neurocomputing*, vol. 312, pp. 135-153, 2018. (<https://doi.org/10.1016/j.neucom.2018.05.083>)

[3] J. Na and W. Hwang, "Deep learning based domain adaptation: A survey," *J. Broadcast Eng.*, vol. 27, no. 4, pp. 511-518, 2022. (<https://doi.org/10.5909/JBE.2022.27.4.511>)

[4] G. Pasqualino, A. Furnari, G. Signorello, and G. M. Farinella, "An unsupervised domain adaptation scheme for single-stage artwork recognition in cultural sites," *Image Vision Comput.*, vol. 107, 104098, 2021. (<https://doi.org/10.48550/arXiv.2008.01882>)

[5] Y. Ganin, E. Ustinova, H. Ajakan, et al., "Domain-adversarial training of neural networks," *The J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2096-2030, 2016.

(<https://doi.org/10.48550/arXiv.1505.07818>)

[6] H. Guan and M. Liu, "Domain adaptation for medical image analysis: A survey," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 3, pp. 1173-1185, 2021. (<https://doi.org/10.1109/TBME.2021.3117407>)

[7] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in *Int. Conf. Mach. Learn. PMLR*, 2015, pp. 1180-1189, 2015. (<https://doi.org/10.48550/arXiv.1409.7495>)

[8] S. Motiian, M. Piccirilli, D. A. Adjeroh, and G. Doretto, "Unified deep supervised domain adaptation and generalization," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 5715-5725, 2017. (<https://doi.org/10.48550/arXiv.1709.10190>)

[9] Y. C. Yu and H. T. Lin, "Semi-supervised domain adaptation with source label adaptation," in *Proc. IEEE/CVF Conf. CVPR*, pp. 24100-24109, 2023. (<https://doi.org/10.48550/arXiv.2302.02335>)

[10] G. Csurka, F. Baradel, B. Chidlovskii, and S. Clinchant, "Discrepancy-based networks for unsupervised domain adaptation: A comparative study," in *Proc. IEEE ICCV*, pp. 2630-2636, 2017. (<https://doi.org/10.1109/iccvw.2017.312>)

[11] M. HassanPour Zonoozi and V. Seydi, "A survey on adversarial domain adaptation," *Neural Process Lett.*, vol. 55, pp. 2429-2469, 2023. (<https://doi.org/10.1007/s11063-022-10977-5>)

[12] K. Saito, D. Kim, S. Sclaroff, and K. Saenko, "Universal domain adaptation through self supervision," *Advances in NIPS*, vol. 33, pp. 16282-16292, 2020. (<https://doi.org/10.48550/arXiv.2002.07953>)

[13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative adversarial nets," *Advances in NIPS*, vol. 27, 2014. (<https://doi.org/10.48550/arXiv.1406.2661>)

[14] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE ICCV*, pp. 2980-

2988, 2017.

(<https://doi.org/10.1109/iccv.2017.324>)

- [15] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, “Feature pyramid networks for object detection, in *Proc. IEEE Conf. CVPR*, pp. 2117-2125, 2017.

(<https://doi.org/10.1109/access.2021.3100369>)

- [16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. CVPR*, pp. 770-778, 2016.

(<https://doi.org/10.1109/cvpr.2016.90>)

- [17] S. A. Orlando, A. Furnari, and G. M. Farinella, “Egocentric visitor localization and artwork detection in cultural sites using synthetic data,” *Pattern Recog. Lett.*, vol. 133, pp. 17-24, 2020.

(<https://doi.org/10.1016/j.patrec.2020.02.014>)

- [18] Z. Zhang and M. Sabuncu, “Generalized cross entropy loss for training deep neural networks with noisy labels,” *Advances in NIPS*, vol. 31, 2018.

(<https://doi.org/10.48550/arXiv.1805.07836>)

- [19] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.

(<https://doi.org/10.48550/arXiv.1412.6980>)

조 함 재 (Hangjae Cho)



2018년 3월~현재 : 인천대학교
임베디드시스템공학과 재학
<관심분야> 인공지능, 컴퓨터
비전

김 정 현 (Jeonghyeon Kim)



2019년 3월~현재 : 인천대학교
임베디드시스템공학과 재학
<관심분야> 인공지능, 컴퓨터
비전

전 경 구 (Kyungkoo Jun)



2001년 5월 : Purdue Univ. 공
학박사 졸업
2004년 3월~현재 : 인천대학교
임베디드시스템공학과 교수
<관심분야> 인공지능, 시스템
소프트웨어, 알고리즘