특집논문 (Special Paper)
방송공학회논문지 제27권 제2호, 2022년 3월 (JBE Vol.27, No.2, March 2022)
https://doi.org/10.5909/JBE.2022.27.2.185
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

객체 검출을 위한 2차원 인조데이터 셋 구축 시스템과 데이터 특징 및 배치 구조에 따른 검출률 분석 : 자동차 번호판 검출을 중점으로

김 상 준[a], 최 진 원[b], 김 도 영[c], 박 구 만[d]‡

2D Artificial Data Set Construction System for Object Detection and Detection Rate Analysis According to Data Characteristics and Arrangement Structure: Focusing on vehicle License Plate Detection

Sang Joon Kim[a], Jin Won Choi[b], Do Young Kim[c], and Gooman Park[d]‡

요 약

최근 객체 인식에 높은 성능을 가진 딥러닝 네트워크가 나오고 있다. 딥러닝을 이용한 객체 인식의 경우 성능 향상을 위해 학습 데이터 셋 구축이 중요하다. 데이터 셋을 구축하기 위해서는 이미지를 수집하고 라벨링 해야 한다. 이 과정은 많은 시간과 인력이 필요하다. 때문에 오픈 데이터 셋을 사용한다. 그러나 방대한 오픈 데이터 셋을 가지고 있지 않는 객체도 존재한다. 그 중 하나가 번호판 검출과 인식에 필요한 데이터이다. 이에 본 논문에서는 이미지를 최소화 하여 대용량 데이터 셋을 만들 수 있는 인조 번호판 생성기 시스템을 제안한다. 또한 인조 번호판 배치구조에 따른 검출률을 분석했다. 분석결과 가장 좋은 배치구조는 FVC_III, B이며 가장 적합한 네트워크는 D2Det이었다. 인조 데이터셋 성능은 실제 데이터셋의 성능보다 2~3%가 낮았지만, 인조 데이터를 구축하는 시간이 실제 데이터셋을 구축하는 시간보다 약 11배 빨라 시간적으로 효율적인 데이터 셋 구축 시스템임을 증명하였다.

Abstract

Recently, deep learning networks with high performance for object recognition are emerging. In the case of object recognition using deep learning, it is important to build a training data set to improve performance. To build a data set, we need to collect and label the images. This process requires a lot of time and manpower. For this reason, open data sets are used. However, there are objects that do not have large open data sets. One of them is data required for license plate detection and recognition. Therefore, in this paper, we propose an artificial license plate generator system that can create large data sets by minimizing images. In addition, the detection rate according to the artificial license plate arrangement structure was analyzed. As a result of the analysis, the best layout structure was FVC_III and B, and the most suitable network was D2Det. Although the artificial data set performance was 2-3% lower than that of the actual data set, the time to build the artificial data was about 11 times faster than the time to build the actual data set, proving that it is a time-efficient data set building system.

Keyword : deep learning, Data labeling, artificial data set, Synthetic data set, Object Detection

Copyright ⓒ 2022 Korean Institute of Broadcast and Media Engineers. All rights reserved.
“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (http://creativecommons.org/licenses/by-nc-nd/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

Ⅰ. 서 론

최근 객체 인식에 높은 성능을 가진 딥러닝 네트워크가 나오고 있다. 딥러닝을 이용한 객체 인식의 경우 성능 향상을 위해 학습 데이터 셋 구축이 중요하다. 데이터 셋을 구축하기 위해서는 이미지를 수집하고 객체가 존재하는 위치를 라벨링 해야 한다. 이 과정은 많은 시간과 인력이 필요하다. 때문에 많은 오픈 데이터 셋을 이용하여 학습을 진행한다[1,2,3]. 그러나 방대한 오픈 데이터 셋이 없는 객체도 존재한다. 번호판 검출과 식별은 방범, 주차 관리, 교통 관리 등과 같은 다양한 환경에 있어 필수 요소로 자리 잡고 있다. 기존 번호판 검출 및 인식 시스템은 영상 처리 기법을 이용했다[4,5]. 최근에는 딥러닝을 활용한 번호판 검출 및 인식 연구가 진행되면서 더욱 발전하고 있다[6,7]. 하지만 번호판 검출을 위한 대용량 데이터 셋이 존재하지 않는다. 따라서 쉽게 대용량 데이터 셋을 만들 수 있는 방법이 필요하다. 이를 위한 연구의 하나로 기존 인조 번호판 데이터 셋을 구축하고 모델 학습을 통해 번호판을 인식하는 연구가 진행되고 있다. T. Björklund 등이 제안한 논문에서는 다양한 조도의 배경 이미지와 인조 번호판을 조합하여 인조 번호판 데이터 셋을 구축하고 이를 학습하였다. 하지만 인조 번호판 생성 시 자동차와 배경의 상관관계를 고려하지 않고 학습하여 다양한 환경에서의 번호판 검출에 문제가 있다[8]. J. Ha 등이 제안한 논문에서는 인조 한국 번호판 이미지를 생성하고 중국 번호판 이미지인 CCPD 데이터 셋과 주행 환경 이미지 BDD 데이터 셋과 결합하여 Mask R-CNN에서 학습하였다. 하지만 인조 번호판의 생성 배치구조가 2개로 한정 적이며 R-CNN 네트워크로만 학습하고, 정확도를 평가하여 인조 번호판의 성능 평가 결과가 명확하다고 보기 어렵다[9]. S. Kim 등이 제안한 논문에서는 인조 번호판, 자동차 범퍼, 배경 이미지를 조합하는 인조 번호판 생성 시스템을 개발하고, 인조 번호판 데이터 셋을 바탕으로 훈련된 모델과 실제 번호판 데이터 셋을 바탕으로 훈련된 모델을 비교하였다. 인조 번호판의 경우 실제 번호판으로 학습하였을 때보다 데이터 셋을 구축하는데 소요되는 시간이 적고, 몇몇 환경에서는 검출률이 더 높게 나왔다. 하지만 번호판의 크기가 작거나 해상도가 낮으면 검출률이 현저히 낮아졌다[10]. 이에 본 논문에서는 자동차가 존재하는 환경의 실제 배경 이미지와 자동차 이미지, 인조 번호판 이미지를 합성하는 인조 번호판 학습 데이터 셋 생성 시스템을 제안한다. 표1은 실제 데이터 수집과 인조 데이터 수집 시간의 차이다. 실제 데이터 수집의 경우 8명의 사람이 인터넷과 직접 자동차를 직접 촬영해 영상을 수집했다. 그리고 수집한 이미지를 라벨링 했다. 인조 데이터 수집의 경우 4명의 사람이 인조 데이터 생성을 위한 데이터 수집과 좌표 정의를 진행했다. 수집 결과 인조 데이터 수집이 실제 데이터를 수집하기 위해 인터넷을 검색하거나 데이터를 촬영하는 것보다 약 11배 빠르며 효율적이었다. 또한 인조 번호판 배치구조에 따른 검출률을 분석하여 가장 좋은 인조 번호판 배치구조와 딥러닝 네트워크를 확인한다. 확인을 위해 5개의 번호판 생성 구조를 바탕으로 10만 장의 이미지를 생성하고 5개의 네트워크에서 학습했다. 검출률 측정을 위해서 서로 다른 7가지 환경을 준비했다.

표 1. 실제 데이터 수집과 인조 데이터 수집 시간 차이
Table 1. The difference between real data collection and artificial data collection time

	Synthetic Data	Real data (Collected from the internet)	Real data (Collected from shooting data)
Image Acquisition	51h	192h	243h
labeling work		371h	
Coordinate definition	26h	-	
Total	77h	806h	

a) 서울과학기술대학교 정보통신미디어공학전공(Dept. of Information Technology and Media Engineering, The graduate School of Nano IT Design Fusion, Seoul National University of Science and Technology)
b) 서울과학기술대학교 기계시스템 디자인공학과(Dept. of Mechanical System Design Engineering, Seoul National University of Science and Technology)
c) 서울과학기술대학교 전기정보공학과(Dept. of Electrical and Information Engineering, Seoul National University of Science and Technology)
d) 서울과학기술대학교 전자 IT 미디어공학과(Dept. of Electronic IT Media Engineering, Seoul National University of Science and Technology)
‡ Corresponding Author : 박구만(Gooman Park)
E-mail: gmpark@seoultech.ac.kr
Tel: +82-2-970-6425
ORCID:https://orcid.org/0000-0002-7055-5568
※ 이 논문의 결과 중 일부는 한국방송·미디어공학회 “2021년 추계학술대회”에서 발표한 바 있음.
※ This work was supported by Institute for Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) in 2022(No. 2017-0-00217, Development of Immersive Signage Based on Variable Transparency and Multiple Layers).
· Manuscript received January 17, 2022; Revised February 14, 2022; Accepted February 14, 2022.

Ⅱ. 관련 연구

1. 딥러닝을 이용한 객체 검출

딥러닝을 이용한 객체 검출은 크게 두 가지 방법으로 나눌 수 있다. 첫 번째는 두 단계(Two-stage)검출이고 두 번째는 한 단계(One-Stage)검출 방법이다. 첫 번째 방법인 두 단계 객체 검출 방법은 위치 제안(Region Proposal)을 사용하여 많은 후보 영역을 선정(Localization)하는 단계와 선정된 후보 영역에서 객체를 분류(Classification)하는 단계로 구성된다. 대표적인 R-CNN은 약 2,000개의 후보 영역을 선정하고 후보 영역을 고정된 크기로 자르고 조정하는 과정을 진행한다. 그리고 다음으로 CNN연산을 수행한다[11]. 이러한 방식은 후보 영역 개수만큼 연산을 진행해야하기 때문에 많은 시간이 소요된다. 두 번째 방법인 한 단계(One-stage) 검출 방법은 객체의 위치를 찾는 문제와 클래스 분류 문제를 동시에 수행하는 방법이다. 대표적인 YOLO(You Lock Only Once)는 객체의 위치와 인식 문제를 회귀 문제로 접근하였다[12]. YOLO의 수행과정은 전체 이미지를 $S \times S$ 그리드로 나누어 그리드 영역에서 클래스 확률과 신뢰도 점수를 측정하여 클래스 점수를 계산하고 bounding box에 위치한 클래스 신뢰도 점수를 출력한다. 본 논문에서는 번호판 검출에 한 단계 검출 방식인 yolo v3[13], SSD[14], M2Det[15], RefineDet[16]과 두 단계 검출 방식인 D2Det[17]을 사용했다.

2. 인조 데이터 셋

딥러닝 네트워크 학습을 위해서 인조 데이터 셋을 구축하는 연구가 진행되고 있다. B. Dwyer 등이 제안한 글에서는 검출하고자 하는 과일 데이터 셋과 구글의 Open Image Dataset을 결합하여 인조 데이터 셋을 구축하고 이를 학습하였다[18]. S. Richter는 대용량 데이터 셋을 만드는데 시간과 비용이 많이 필요하기 때문에 다음과 같은 방법을 제안했다. 현실적인 오픈월드의 컴퓨터 게임에서 2만 5천장의 이미지를 추출하고 답안지를 생성하여 학습하는 방법이다. 이 방법은 CamVid 데이터 셋[20]보다 높은 성능을 냈다[19]. SYNTHIA는 운전 시나리오 맥락을 돕기 위해 생성된 인조 데이터 셋이다. 가상 도시에서 렌더링된 현실적인 프레임 모음으로 구성되어 있다. 하늘, 건물, 도로, 보도 등 13개의 클래스를 가지고 있다[21]. M. Müller는 UAV(unmanned aerial vehicle)가 새로운 환경에서도 손쉽게 비행 학습을 할 수 있는 가상의 공간을 만들어 UAV를 손쉽게 학습하는 방법을 제안했다[22]. NAVER LABS에서는 자율 주행에 필요한 데이터와 학습을 위해 유니티 엔진을 이용해 가상의 공간을 만들어 학습하고 시험하는 방법을 제안했다[23]. OpenAI는 인조 데이터를 제작하는 시간, 비용, 인력이 필요하다는 단점을 보완하기 위해 저비용으로 대량의 이미지를 합성하여 데이터 셋을 만들고 정확도를 향상시키는 Domain Randomization기법을 객체 검출에 적용하는 방법을 제안했다. Domain Randomization는 찾고자 하는 객체의 3D 모델과 임의의 배경, 텍스처 이미지들이 준비물이 된다. 이 때 임의의 배경 이미지는 실제 환경과 유사한 배경이면 좋지만 그렇지 않은 배경을 이용해도 무방하다. 이러한 방식으로 임의의 배경에 임의의 텍스처를 삽입하고, source item을 임의의 geometric shape를 이용하여 임의의 위치에 합성하면 객체 검출 학습에 필요한 이미지와 라벨(Label)을 쉽게 제작할 수 있다[24,25].

Ⅲ. 인조 번호판 데이터 생성 시스템

1. 이미지 배치 구조

본 논문에서는 실험하고자 하는 인조 번호판 이미지 배치 구조는 표 2와 같다. CB의 배치 구조는 자동차 범퍼와 인조 번호판이 합성된 배치 구조이다. CB의 약자는 Car Bumper이다. 자동차 범퍼를 중앙에 배치하고 미리 정의된 번호판 좌표에 인조 번호판을 배치한다. 그리고 크기를 랜덤으로 변화시킨 후 배경 이미지와 결합한다. 미리 정의된 번호판 좌표는 그림 1과 같이 이미지에서 번호판 영역이라고 추정되는 4개의 꼭지 점으로 이루어진 데이터이다. 이 데이터는 선행 작업을 통해 미리 만들어진다. FVC_I의 배치 구조는 자동차 이미지와 인조 번호판이 합성된 배치 구조이다. FVC의 약자는 Full view of the car이다. CB와 같은 배치구조를 가지고 있으나 자동차의 크기는 변하지 않

표 2. 인조 번호판 이미지 배치 구조
Table 2. Artificial Number Plate Image Layout Structure

No.1	1	2	3
Type	CB (Car Bumper)	FVC_I (full view of the car)	FVC_II
Image			
explanation	Combined the bumper and artificial license plate	Combined entire image of the car and artificial license plate	Combined of the entire image of a car with a size variation and an artificial license plate
No.1	4	5	6
구분	FVC_III	CI (Car Image)	RI (Real Image)
이미지			
explanation	It is an artificial license plate combined with the overall image of a car that varies in size and number of vehicles.	Combined real car image and artificial license plate	Real car image

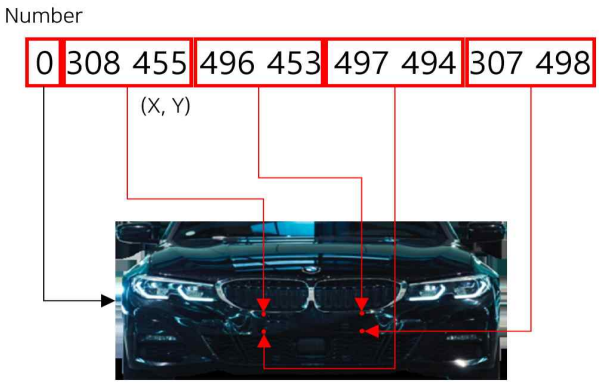

그림 1. 번호판 영역으로 추정되는 영역의 4개의 좌표
Fig. 1. The four coordinates of the area that is assumed to be the license plate area

는다. FVC_II의 배치 구조는 FVC_I과 같은 배치 구조를 가지고 있으나 자동차의 크기는 랜덤으로 변한다. FVC_III은 1~10까지 랜덤으로 적용되는 배치 구조에 따라 여러 자동차 이미지를 배치하고 자동차의 정의된 위치에 인조 번호판을 배치한다. 그리고 크기를 랜덤으로 변화시킨 후 배경 이미지와 결합된다. CI의 배치 구조는 자동차를 찍은 이미지에 미리 정의된 번호판 좌표에 인조 번호판을 배치하는 형태이다. CI의 약자는 Car Image이다. 여기서 실제 이미지와 CI의 이미지가 다른 점은 같은 이미지에 여러 번호판을 학습할 수 있는 것이다. 그림 1에서 가장 앞에 있는 숫자는 이미지에서 범퍼의 고유 번호이다. 두 번째 숫자부터는 4개 꼭지 점 x, y 좌표이다. 첫 번째 x, y 좌표는 번호판의 좌측 상단의 꼭지 점 좌표이다. 두 번째 x, y좌표는 우측 상단 꼭지 점 좌표이다. 세 번째 x, y좌표는 좌측 하단의 꼭지 점 좌표이다. 네 번째 x, y좌표는 우측 하단의 꼭지 점 좌표이다.

2. 제안된 시스템 개요

본 논문에서 제안하는 인조 번호판 학습 데이터 생성 시스템의 전체적인 과정은 그림 2와 같이 5가지 모듈로 구성된다. 5가지 모듈은 인조 번호판 생성 모듈, 잡음 생성 모듈, 번호판 좌표 변환 모듈, 이미지-번호판 결합 모듈, 배경 이미지 모듈이다. 첫 번째 인조 번호판 생성을 위한 모듈은

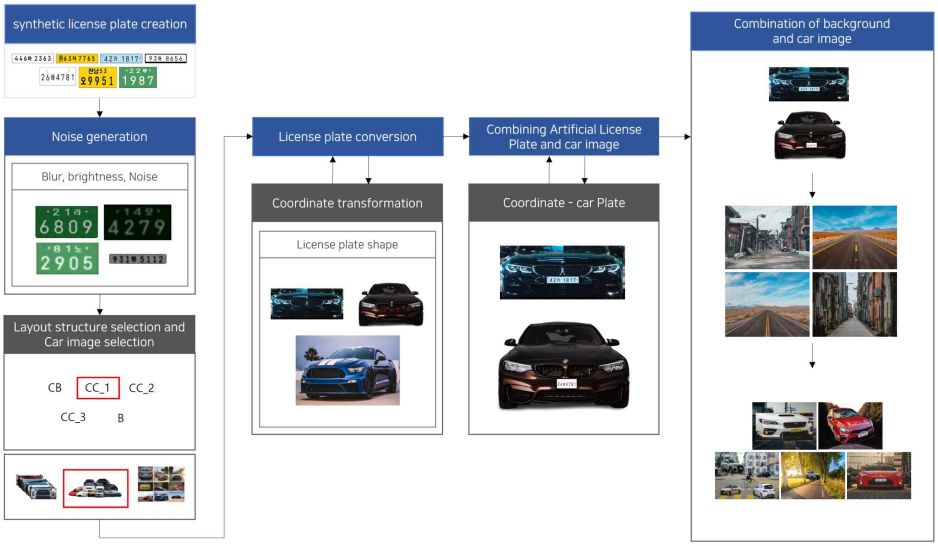

그림 2. 제안된 시스템 구성과 모듈
Fig. 2. Proposed system configuration and modules

번호판 생성에 필요한 이미지를 이용하여 인조 번호판을 생성 한다. 두 번째 잡음 생성 모듈은 명도, 채도, 블러, 노이즈를 이용하여 생성된 번호판에 잡음을 추가한다. 세 번째 번호판 좌표 변환 모듈은 생성된 번호판의 좌표를 미리 정의된 좌표로 변환해준다. 네 번째 이미지-번호판 결합 모듈은 좌표 변환된 번호판과 자동차를 결합해준다. 다섯 번째 배경 이미지 모듈은 번호판과 결합된 자동차 이미지와 배경 이미지를 결합해준다.

3. 인조 번호판 생성

인조 번호판 생성을 위해서 그림 3과 같이 번호판 생성에 필요한 번호판, 심볼, 볼트, 숫자, 글자, 지역 이미지가 필요하다. 이미지들을 이용해 원하는 양의 번호판을 생성한다. 여기서 번호판은 국토교통부에서 제공하는 공식 번호판 규격과 문자 형태에 맞춰 흰색, 노란색, 초록색, 전기차 번호판을 구분하고 다시 번호판 종횡비와 문자의 배치에 따라

그림 3. 인조 번호판 생성을 위한 구성 요소
Fig. 3. Components for creating artificial license plates

표 3. 학습 이미지를 만들기 위한 합성 이미지
Table 3. Composite images to create training images

Background images	Car bumper images
Image of cars	Images of each car

색상별로 두 가지 형태의 번호판 형태로 나누어 7종류 번호판 형태를 이용했다. 인조 번호판 생성은 다음과 같이 진행된다. 만들고자 하는 번호판이 임의로 선택된다. 번호판의 종류는 자가용(일반) 2자리 배치, 자가용(일반) 3자리 배치, 자가용(대형), 전기차, 사업용(보통), 사업용(대형)이다. 번호판을 선택하면 번호판 내에 위치할 문자 총 84개 중 임의로 문자가 선택되어 번호판에 배치된다. 볼트의 경우 번호판 내에 임의로 선택돼 배치된다. 심볼은 전기차에 해당되면 배치된다.

4. 배치구조에 따른 합성 이미지 선택

인조 번호판을 생성하고 합성 이미지들의 배치구조를 선택한다. 배치 구조에 따라 합성 이미지가 다르기 때문에 배치 구조에 해당하는 이미지를 준비해야 한다. 본 논문에서는 표3과 같이 배경 이미지 총150개, 자동차 범퍼 이미지 150개, 자동차 이미지 152개, 자동차를 찍은 이미지 5000장을 준비해 사용했다. 배경 이미지는 도로가 포함된 도시 이미지 50장, 도로가 포함된 풍경 이미지 50장, 도로가 포함된 골목 이미지 50장으로 자동차와 밀접한 관련이 있는 이미지로 준비했다. 자동차 범퍼와 자동차 이미지는 번호판이 없는 현대, 기아, 테슬라, 포드, BMW, 아우디, 벤츠 외 11개의 제조사 이미지를 준비했다. 차종은 소형7개, 중형10개, 대형7개, SUV 15개, 스포츠카 8개로 총 47개를 준비했다. 자동차를 찍은 이미지는 상업적, 비상업적 사용이 가능한 인터넷 사이트와 촬영 영상을 사용했다.

Ⅳ. 실험 결과

본 실험에서는 생성된 인조 번호판 데이터 셋의 성능을 여러 가지 방면으로 평가하기 위해 YOLO v3[13], SSD[14], M2Det[15], RefineDet[16], D2Det[17] 네트워크를 이용했다. 성능 평가를 위해서 그림 4와 같이 7가지 다양한 실험환경에서 촬영된 동영상을 이용했다. 그림 4에서 첫 번째, 두 번째 실험환경은 주차된 자동차의 번호판을 가까이 차례대로 찍은 영상이다. 세 번째, 네 번째, 다섯 번째 실험환경은 자동차 주행 중 번호판을 찍은 영상이다. 여섯 번째, 일곱 번째 실험환경은 CCTV가 자동차를 찍는 거리와 각도에서 찍은 영상이다.

인조 번호판 검출 결과는 그림 5, 6, 7, 8, 9, 10, 11과

그림 4. 인조 번호판 데이터 셋 성능 평가를 위한 실험환경
Fig. 4. Experimental environment for performance evaluation of artificial license plate data set

그림 5. Yolo v3 네트워크를 이용한 Video1 검출 결과
Fig. 5. Video1 detection result using Yolo v3 network

그림 6. Yolo v3 네트워크를 이용한 Video2 검출 결과
Fig. 6. Video2 detection result using Yolo v3 network

그림 7. Yolo v3 네트워크를 이용한 Video3 검출 결과
Fig. 7. Video3 detection result using Yolo v3 network

그림 8. Yolo v3 네트워크를 이용한 Video4 검출 결과
Fig 8. Video4 detection result using Yolo v3 network

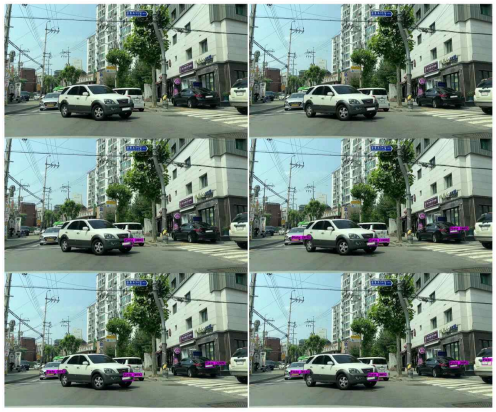

그림 9. Yolo v3 네트워크를 이용한 Video5 검출 결과
Fig. 9. Video5 detection result using Yolo v3 network

그림 10. Yolo v3 네트워크를 이용한 Video6 검출 결과
Fig. 10. Video6 detection result using Yolo v3 network

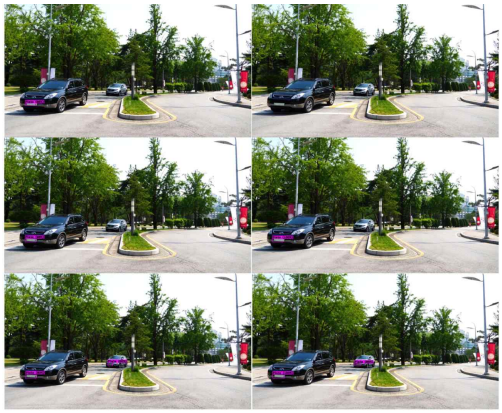

그림 11. Yolo v3 네트워크를 이용한 Video7 검출 결과
Fig. 11. Video7 detection result using Yolo v3 network

같다. 그림 5, 6, 7, 8, 9, 10, 11은 Yolo v3 네트워크의 검출 결과로 첫 번째 열 좌측 배치구조 CB의 검출결과이다. 첫 번째 열 우측 배치구조는 FVC_I의 검출결과이다. 두 번째 열 좌측 배치구조는 FVC_II의 검출결과이다. 두 번째 열 우측 배치구조는 FVC_III의 검출결과이다. 세 번째 열 좌측 배치구조는 CI의 검출결과이다. 세 번째 열 우측 배치구조는 RI의 검출결과이다. 각 네트워크의 배치에 따른 검출률은 표 4와 같다. 전체적으로 가장 안 좋은 인식률을 보여준 배치 방법은 자동차 범퍼와 결합한 배치구조 CB이다. 특히 CCTV가 자동차를 찍는 거리와 각도에서 찍는 영상인 Video6, Video7에서는 인식률이 10%를 넘지 못했다. 이는 학습 이미지 특성 상 크기가 큰 범퍼와 인조 번호판을 학습

표 4. 인조 번호판 데이터를 이용한 검출 결과
Table 4. Detection results using artificial license plate data

video	layout structure	Method (mAP 50)				
		yolo v3	SSD	D2Det	M2Det	RefineDet
Video_1	CB	37.76	48.2	54.3	35.57	48.05
	FVC_I	82.73	74.8	64.8	51.6	80.74
	FVC_II	67.43	71.3	80.5	68.18	87.61
	FVC_III	79.54	70.7	87.2	80.10	81.30
	CI	88.22	85.2	89.9	86.86	86.22
	RI	89.19	87.69	92.8	88.59	87.85
Video_2	CB	52.28	50.9	61.9	53.79	36.30
	FVC_I	89.91	77.0	70.5	69.4	90.39
	FVC_II	87.66	78.4	79.7	82.70	90.79
	FVC_III	89.07	71.4	89.6	91.95	92.63
	CI	93.75	91.7	82.6	95.02	84.87
	RI	93.45	92.88	88.72	96.44	91.65
Video_3	CB	7	36.1	10.8	1.30	59.14
	FVC_I	31.34	24.9	5.4	2.6	38.58
	FVC_II	15.22	23.2	63	26.83	44.08
	FVC_III	51.19	27.3	79.5	37.61	65.05
	CI	84.7	65.2	87.1	63.54	80.54
	RI	82.6	71.66	86.45	69.87	85.42
Video_4	CB	15.51	29.8	11.6	2.6	44.28
	FVC_I	30.47	16.4	4.4	6.3	36.63
	FVC_II	19.51	21.5	41.4	28.48	36.29
	FVC_III	40.08	31.3	67.9	48.1	59.01
	CI	72.2	59.1	68.0	41.15	65.69
	RI	74.1	62.4	70.56	52.14	70.12
Video_5	CB	11.82	48.6	12.2	1.43	68.28
	FVC_I	46.53	40.5	6.1	8.2	62.7
	FVC_II	27.5	33.3	65.68	48.32	65.62
	FVC_III	55.42	37.4	85.4	64.59	71.02
	CI	76.08	65.7	87.2	68.46	81.9
	RI	75.25	67.55	90.1	70.34	84.62
Video_6	CB	3.01	29.7	1	0	46.12
	FVC_I	8.2	3.6	0	0	27.5
	FVC_II	16.42	8.6	46.8	38.14	45.07
	FVC_III	69.98	49.1	89.5	54.25	74.11
	CI	97.67	68.2	83.7	33.85	74.05
	RI	96.48	74.81	84.55	31.48	75.61
Video_7	CB	9.53	32.5	3.9	2.47	41.1
	FVC_I	27.35	24.1	1.0	4.1	32.68
	FVC_II	23.35	16.3	43.2	44.71	32.31
	FVC_III	66.13	35.0	85.9	69.51	69.64
	CI	88.04	63.8	83.9	76.55	73.36
	RI	89.57	64.15	88.04	78.24	77.21

했기 때문이다. 또한 한 장의 이미지에 한 개의 범퍼만 존재하기 때문에 다수의 자동차 이미지가 나오는 Video3, Video4, Video5에서도 좋은 인식률을 보여주지 못했다. 자동차의 번호판을 가까이 차례로 찍은 Video_2에서는 다른 Video보다 좋은 성능을 보여준다. CB 배치 구조에서 가장 좋은 성능을 보인 네트워크는 RefineDet 네트워크이다. FVC_I의 구조는 Video 1, Video2에서 무난한 모습을 보여줬다. 하지만 작은 번호판이 섞여있는 Video3, Video4, Video5그리고 작은 번호판이 많은 Video6, Video7에서는 낮은 정확도를 보였다. 이는 자동차 전체적인 부분이 학습되었지만 번호판의 크기가 일정하고 작은 번호판 데이터 학습 정보 획득에 제한이 있었기 때문이다. FVC_I의 구조

에서 가장 좋은 성능을 보인 네트워크는 RefineDet 네트워크이다. FVC_II의 구조는 전체적으로 FVC_I, FVC_III의 구조보다 낮은 정확도를 보여준다. 이는 자동차와 번호판 데이터 크기 분산으로 크기별 학습 데이터 부족으로 보인다. 또한 CB와 마찬가지로 한 장의 이미지에 한 개의 자동차만 존재하기 때문에 다수의 자동차 이미지가 나오는 Video에서도 낮은 정확도를 보인다. FVC_II의 구조에서 가장 좋은 성능을 보인 네트워크는 D2Det 네트워크이다. FVC_III의 구조는 전체적으로 모든 비디오에서 평균에 해당한다. 이는 FVC_II와는 다르게 여러 대의 자동차가 한 이미지에 담겨져 있고 모든 자동차의 크기가 골고루 분포되었기 때문이다. 가장 좋은 네트워크는 D2Det 네트워크이다.

CI의 구조는 인조 데이터 배치 구조에서 가장 좋은 성능을 보여준다. 또한 몇몇 환경에서는 실제 이미지로 학습한 데이터 보다 더 좋은 인식률을 보여줬다. 이는 자연스러운 환경 배치 구조와 다른 번호판 구성들을 한 장의 이미지에서 재활용 할 수 있기 때문으로 보인다. CI의 구조상 가장 좋은 네트워크는 yolo v3 네트워크이다. D2Det 네트워크는 2단계 검출로 모든 배치구조에서 좋은 성능을 보여줬다. 특히 FVC_II, FVC_III에서 다른 네트워크보다 좋은 성능을 보여줬다. 이는 네트워크 특성상 배경 영역의 영향을 줄이고 정확하게 현지화 하여 오브젝트를 검출하기 때문으로 보인다. 하지만 fps는 3fps를 기록하여 매우 느린 속도를 보여준다. RefinDet은 몇몇 상황에서 좋은 성능을 보여줬다. 특히 video_2 FVC배치구조에서 가장 좋은 성능을 보여줬다. 이는 위치와 크기를 정교하게 예측할 수 있는 anchor box와 high-level feature map의 정보를 활용했기 때문에 형상을 이해하는 부분에서 높은 인식률을 보여준 것으로 보인다. 그러나 정상적인 데이터가 주어졌을 때는 다른 네트워크에 비해 낮은 인식률을 보여준다. yolo v3는 작은 객체 혹은 오브젝트의 형상이 전체적으로 표현이 되지 않으면 학습에 무리가 있어 보인다. 하지만 전체적인 형상을 학습해주면 다른 네트워크보다 좋은 검출율을 보여줬다.

제안 방법과 기존방법의 비교는 표 5와 같다. 비교를 위한 네트워크는 모든 배치구조에서 좋은 성능을 보여준 D2Det을 이용했다. 기존 방법의 학습 파일과 인조 데이터는 찾을 수 없어 논문을 바탕으로 데이터를 만들었다. 가장 좋은 성능을 보여준 배치 방법은 본 논문에서 제안한 CI배치 방법이다. T. Björklund 등이 제안한 방법은 번호판과 밀접한 관련이 없는 배경으로 학습되어 전체적으로 가장 안 좋은 검출률을 보여준다. J. Ha 등이 제안한 방법 중 BDD 방법의 경우 T. Björklund 등이 제안한 방법과 비슷하지만 번호판과 관련있는 배경으로 학습되어 T. Björklund 등이 제안한 방법보다는 높은 성능을 보였다. CCPD의 방법의 경우 본 논문에서 제안하는 CI 방법과 동일하지만 자동차의 번호판 위치와 배치가 정면 위주의 데이터로 다양성이 없어 CI 방법보다는 낮은 검출률을 보였다.

V. 결 론

최근 객체 인식에 높은 성능을 가진 딥러닝 네트워크가 나오고 있다. 딥러닝을 이용한 객체 인식의 경우 성능 향상을 위해 학습 데이터 셋 구축이 중요하다. 데이터 셋을 구축하기 위해서는 이미지를 수집하고 객체가 존재하는 위치를

표 5. 제안 방법과 기존 방법 비교
Table 5. Comparison of the proposed method with the existing method

Method	layout structure	Video						
		Video_1	Video_2	Video_3	Video_4	Video_5	Video_6	Video_7
D2Det	T. Björklund	28.4	33.8	8.4	7.25	10.92	0	1
	J. Ha(CCPD)	74.2	80.1	75.1	56.4	83	78.4	78.2
	J. Ha(BDD)	36	47.2	10.14	9.8	15.1	0	1
	S. Kim	54.3	61.9	10.8	11.6	12.2	1	3.9
	Our(FVC_III)	87.2	89.6	79.5	67.9	85.4	89.5	85.9
	Our(CI)	89.9	82.6	87.1	68.0	87.2	83.7	83.9

라벨링 해야 한다. 이 과정은 많은 시간과 인력이 필요하다. 때문에 많은 오픈 데이터 셋을 이용하여 학습을 진행한다. 그러나 방대한 오픈 데이터 셋을 가지고 있지 않는 데이터 셋도 존재한다. 그 중 하나가 번호판 검출과 인식에 필요한 데이터이다. 이에 본 논문에서는 이미지를 최소화 하여 대용량 학습 데이터 셋을 만들 수 있는 인조 번호판 생성기 시스템을 제안하였다. 또한 인조 번호판 배치구조에 따른 검출률을 분석하여 가장 좋은 인조 번호판 배치구조를 제안하였다. 효용성 증명을 위해 5개의 번호판 생성 구조를 바탕으로 10만 장의 이미지를 생성하고 5개의 네트워크에서 학습하였다. 정확도 측정을 위해 서로 다른 7가지 환경을 준비했다. 실험 결과 실제 데이터로 구축한 데이터셋으로 학습한 네트워크의 AP50에 비해 2~3% 낮았다. 하지만 표 1에서 알 수 있듯이 인조 데이터 셋 구축 시스템으로 데이터를 생성하는 것이 실제 데이터를 수집하는 라벨링하는 작업보다 시간을 약 11배 단축할 수 있어 시간적 측면에서 훨씬 효율적이었다. 향후 연구로 배경에 적절하게 이미지가 배치될 수 있는 네트워크를 고안하여 실제 데이터와 비슷한 검출율을 보이지만 쉽고 빠르게 제작할 수 있는 데이터 생성기 시스템을 연구할 예정이다.

참 고 문 헌 (References)

[1] OpenImage [Internet] https://storage.googleapis.com/openimages/web/ factsfigures.html

[2] MS CoCO [Internet] https://cocodataset.org/#home

[3] CIFAT-10 [Internet] http://www.cs.toronto.edu/~kriz/cifar.html

[4] Jin-Ho Kim, and Duck-soo Noh "Vehicle License Plate Recognition System By Edge-based Segment Image Generation", JOURNAL OF THE KOREA CONTENTS ASSOCIATION 12(3), 2012.3, 9-16 doi: https://doi.org/10.5392/JKCA.2012.12.03.009

[5] Hyoung-chul oh and Jong-Ho choi "A Recognition Algorithm of Car License Plate using Edge Projection and Directivity Vector", The Journal of Korean Institute of Information Technology 7(1), 2009.2, 83-92. https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE01149744

[6] Seung-ju Lee and Goo-man Park, "Proposal for License Plate Recognition Using Synthetic Data and Vehicle Type Recognition System", JOURNAL OF BROADCAST ENGINEERING 25(5), 2020.9, 776-788. doi: https://doi.org/10.5909/JBE.2020.25.5.776

[7] Jung-Hwan Kim, Joon-Hong Lim, "License Plate Detection and Recognition Algorithm using Deep Learning", journal of IKEEE 23(2), 2019.6, 642-651. doi: http://dx.doi.org/10.7471/ikeee.2019.23.2.642

[8] T. Björklund, A. Fiandrotti, M. Annarumma, G. Francini and E. Magli, “Automatic License Plate Recognition with Convolutional Neural Networks Trained on Synthetic Data , “ International Workshop on Multimedia Signal Processing (MMSP), pp. 1-6, 2017. doi: http://dx.doi.org/10.1109/MMSP.2017.8122260

[9] J. Ha "Detection of Korea License Plate by Mask R-CNN using Composite Image" Journal of Institute of Control, Robotics and Systems 26(9), 2020.9, 778-783. doi: http://dx.doi.org/10.5302/J.ICROS.2020.20.0070

[10] S. Kim, Y. Lee and G. Park, “Artificial license plate generation system for vehicle license plate detection“ Korean society internet Information, VOL 21 NO.01 PP.33~34. 2020. https://www.eiric.or.kr/literature/ser_view.php? SnxGubun=IN KO&mode=total&searchCate=literature&gu=INME020E0&cmd=qryview&SnxIndxNum=232898&rownum=&totalCnt=2&rownum=2&q1_t=7J247KGwIOuyiO2YuO2MkA==&listUrl=L3NlYXJjaC9yZXN1bHQucGhwP1NueEd1YnVuPUlOS08mbW9kZT10b3RhbCZzZWFyY2hDYXRlPWxpdGVyYXR1cmUmcTE9JUMwJUNFJUMxJUI2KyVCOSVGOCVDOCVBMyVDNiVDNyZ4PTAmeT0w&q1=%C0%CE%C1%B6+%B9%F8%C8%A3%C6%C7&kci=

[11] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. arXiv:1311.2524 (2014). doi: https://doi.org/10.48550/arXiv.1311.2524

[12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640, 2015. doi: https://doi.org/10.48550/arXiv.1506.02640

[13] Joseph Redmon and Ali Farhadi. YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018. doi: https://doi.org/10.48550/arXiv.1804.02767

[14] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed. Ssd: Single shot multibox detector. arXiv preprint arXiv:1512.02325, 2015. doi: https://doi.org/10.1007/978-3-319-46448-0_2

[15] Qijie Zhao, Tao Sheng, Yongtao Wang, Zhi Tang, Ying Chen, Ling Cai, and Haibin Ling. M2det: A single-shot object detector based on multi-level feature pyramid network. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), volume 33, pages 9259–9266, 2019. doi: https://doi.org/10.1609/aaai.v33i01.33019259

[16] Zhang, S., Wen, L., Bian, X., Lei, Z., Li, S.Z.: Single-shot refinement neural network for object detection. arXiv preprint arXiv:1711.06897 doi: https://doi.org/10.48550/arXiv.1711.06897

[17] J. Cao, H. Cholakkal, R. M. Anwer, F. S. Khan, Y. Pang and L. Shao, "D2Det: Towards High Quality Object Detection and Instance Segmentation," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11482-11491, 2020.

doi: https://doi.org/10.1109/CVPR42600.2020.01150
[18] How to Create a synthetic Dataset for computer vision [Internet] https://blog.roboflow.com/how-to-create-a-synthetic-dataset-for-computer-vision/
[19] S.R Richter, V. Vineet, S. Roth, and V. Koltun “Playing for Data: Ground Truth from Computer Games.” European Conference on Computer Vision – ECCV 2016 pp 102-118, 2016.
doi: https://doi.org/10.48550/arXiv.1608.02192
[20] CamVid Database [Internet] http://mi.eng.cam.ac.uk/research/projects/ VideoRec/
[21] SYNTHIA dataset [Internet] http://synthia-dataset.net/
[22] M. Müller, V. Casser, N. Smith, D.L Michels and B. Ghanem “Teaching UAVs to Race: End-to-End Regression of Agile Controls in Simulation” European Conference on Computer Vision – ECCV 2018 pp 11-29.
doi: 2018. https://doi.org/10.48550/arXiv.1708.05884
[23] NAVER LABS KITTI dataset [Internet] https://europe.naverlabs.com/Research/Computer-Vision/Proxy-Virtual -Worlds/
[24] J. Tremblay, et. al. "Training Deep Networks with Synthetic Data: Bridging the Reality Gap by Domain Randomization". Conference on Computer Vision and Pattern Recognition Workshops(CVPRW) 2018.
doi: https://doi.org/10.48550/arXiv.1804.06516
[25] Training Deep Networks with Synthetic Data Bridging the Reality Gap by Domain Randomization Review [Internet] https://hoya012.github.io/blog/Tutorials-of-Object-Detection-Using-Deep-Learning-performance-o ne/

저 자 소 개

김 상 준

- 2017년 3월 ~ 2019년 2월 : 서울미디어대학원대학교 미디어공학전공 공학석사
- 2019년 3월 ~ 현재 : 서울과학기술대학교 정보통신미디어공학 공학박사
- ORCID : https://orcid.org/0000-0001-7498-6149
- 주관심분야 : 인공지능, 컴퓨터그래픽스, 증강현실, 프로젝션 맵핑

최 진 원

- 2016년 3월 ~ 현재 : 서울과학기술대학교 기계시스템 디자인공학과 공학사
- ORCID : https://orcid.org/0000-0003-1349-337X
- 주관심분야 : 인공지능, 컴퓨터비전

김 도 영

- 2016년 3월 ~ 현재 : 서울과학기술대학교 전기정보공학과 공학사
- ORCID : https://orcid.org/0000-0002-7158-7238
- 주관심분야 : 인공지능, 컴퓨터비전

저 자 소 개

박 구 만

- 1984년 : 한국항공대학교 전자공학과 공학사
- 1986년 : 연세대학교 대학원 전자공학과 석사
- 1991년 : 연세대학교 대학원 전자공학과 박사
- 1991년 ~ 1996년 : 삼성전자 신호처리연구소 선임연구원
- 1999년 ~ 현재 : 서울과학기술대학교 전자IT미디어공학과 교수
- 2006년 ~ 2007년 : Georgia Institute of Technology, Dept of ECE. Visiting Scholar
- 2016년 ~ 2017년 : 서울과학기술대학교 나노 IT디자인융합대학원 원장
- ORCID : https://orcid.org/0000-0002-7055-5568
- 주관심분야 : 컴퓨터비전, 실감미디어