KIPS Tr. Software and Data Eng.
Vol.5, No.11 pp.511~520 pISSN: 2287-5905

http://dx.doi.org/10.3745/KTSDE.2016.5.11.511

Real-Time License Plate Detection Based on Faster R-CNN

Dongsuk Lee[†] · Sook Yoon[††] · Jaehwan Lee[†††] · Dong Sun Park[††††]

ABSTRACT

Automatic License Plate Detection (ALPD) is a key technology for a efficient traffic control. It is used to improve work efficiency in many applications such as toll payment systems and parking and traffic management. Until recently, the hand-crafted features made for image processing are used to detect license plates in most studies. It has the advantage in speed. but can degrade the detection rate with respect to various environmental changes. In this paper, we propose a way to utilize a Faster Region based Convolutional Neural Networks (Faster R-CNN) and a Conventional Convolutional Neural Networks (CNN), which improves the computational speed and is robust against changed environments. The module based on Faster R-CNN is used to detect license plate candidate regions from images and is followed by the module based on CNN to remove False Positives from the candidates. As a result, we achieved a detection rate of 99.94% from images captured under various environments. In addition, the average operating speed is 80ms/image. We implemented a fast and robust Real-Time License Plate Detection System.

Keywords : License Plate, Deep Learning, Convolutional Neural Network, Faster Region Based Convolutional Neural Network

Faster R-CNN 기반의 실시간 번호판 검출

이 동 석[†] · 윤 숙[††] · 이 재 환[†††] · 박 동 선[††††]

요 약

자동차 번호판 검출 자동화(ALPD: Automatic License Plate Detection) 시스템은 효율적인 교통 관제를 위한 핵심 기술이며, 통행료 지불 시스템, 주차장 및 교통 관리와 같은 많은 응용에 사용되어 업무의 효율을 높이고 있다. 최근까지의 ALPD에 관한 연구에서는 주로 영상처리를 위해 설계된 기존의 특징들을 추출하여 번호판 검출에 사용해왔다. 이러한 종래의 방법은 속도에 이점은 있으나, 다양한 환경 변화에 따른 성능 저하를 보였다. 본 논문에서는 전반적인 성능을 향상시키기 위하여 Faster R-CNN과 CNN으로 구성되는 두 단 구조를 활용하는 방법을 제안한다. 이를 통해 동작 속도를 향상시키고, 다양한 환경변화에 강인하도록 구성하였다. 첫 번째 단계에서는 Faster R-CNN을 적용하여 번호판 영역 후보영역들을 선별하며, 두 번째 단에서 CNN을 활용하여 후보영역들 중에서 False Positives를 제거함으로써 검출률을 향상시켰다. 이를 통해 ZFNet을 기반으로 하여 99.94%의 검출률을 달성하였다. 또한 평균 운용시간은 80ms/image로써 빠르고 강인한 실시간 번호판 검출 시스템을 구현할 수 있었다.

키워드 : 자동차 번호판, Deep Learning, 컨볼루션 신경망, Faster Region Based Convolutional Neural Networks

1. 서 론

최근 차량의 증가와 교통 환경의 변화들로 인하여, 좀 더 효율적인 교통 관제 시스템에 대한 수요가 늘어나고 있다. 특히 자동화된 번호판 검출 시스템은 지능형 교통 시스템의 핵심적인 요소이며, 이미 상용화 되어 통행료 지불 시스템, 주차장 및 교통 관리와 같은 다양한 분야에 응용되어 사람의 단순 반복 업무를 대신함으로서 전반적인 업무의 효율성을 높이고 있다. '번호판 검출'의 성능에는 Table 1과 같은 번호판 및 환경의 변화가 많은 영향을 준다. 카메라 센서가 환경에 대한 민감성을 줄이고, 해상도를 개선하는 방향으로 발전하고 있으며, 영상정보 외에 적외선 센서를 병행해 사용하여 이러한 문제들을 줄이고는 있으나, 이는 시스템의 구축 비용의 상승등으로 이어지기 때문에 환경 변화에 강인하고 높은 검출 성능을 위한 연구는 지속적으로 이루어지고 있다. 일반적인 번호판 검출 시스템들은 '특징 추출', '탐색', '분류', '위치 추정'과 같은 단계들로 이루어지며 각 단계들에는 개별적인 알고리즘들이 적용되었다. 특히 특징 추출과정에서는 Edge, Conner, Haar-like, HOG(Histogram of Oriented

※ 본 성과물은(논문, 산업재산권, 품종보호권 등)은 농촌진흥청 연구사업 (세부과제번호: PJ0120642016)의 지원에 의해 이루어진 것임.
† 준 회 원 : 전북대학교 IT 융합연구센터 연구원
†† 비 회 원 : 목포대학교 멀티미디어공학과 부교수
††† 비 회 원 : 전북대학교 전자정보공학부 박사과정
†††† 정 회 원 : 전북대학교 전자공학부 교수
Manuscript Received : October 4, 2016
Accepted : October 12, 2016
* Corresponding Author : Sook Yoon(syoon@mokpo.ac.kr)

Gradient)와 같이 전문가들에 의해 설계된 특징들을 사용해 왔다. 이러한 방법들은 Table 1과 같은 요인들에 대하여 민감하여, 입력 영상을 그대로 활용하는데 어려움을 주었다. 이러한 문제의 해결을 위하여 이진화, 평활화와 같은 다양한 전처리를 적용하게 되었고, 이것은 또 다른 잡음들의 원인이 되기도 하였다. 최근까지도 이러한 문제들로 인하여 성능 향상에 많은 어려움들이 있어왔다.

본 논문에서는 Faster R-CNN[14]과 CNN[15, 20]을 적용함으로써, 다양한 환경 변화에 강인하고, 높은 정확성을 지닌 실시간 번호판 검출 시스템을 제안한다. 제안하는 시스템의 번호판 검출과정은 '검출'과 '필터링'의 두 단으로 이루어져 있으며, '검출' 단계에서 Faster R-CNN[14]를 적용하여 입력과 번호판의 크기에 상관없이 네트워크에서 직접 번호판의 영역을 추정하며, 이와 함께 연산량의 이득을 얻고, '필터' 단계에서는 기존의 분류기보다 상대적으로 뛰어난 일반화 능력을 지닌 CNN[5]를 적용함으로서 충분하지 못한 훈련샘플로 인하여 발생할 수 있는 False Positive를 효과적으로 제거할 수 있도록 구성하였다. 논문의 구성은 2장에서 기존의 번호판 검출 방법들에 대해서 이야기하고, Faster R-CNN[14]과 CNN[15, 20]을 소개하며, 3장에서는 제안하는 번호판 검출 시스템을 설명한다. 이후 4장에서는 시스템의 훈련 및 성능평가를 실시하며, 5장에서 결론을 맺는다.

Table 1. Variation Factors of License Plates

	Factors	Explanation
Plate variations	Location	location variations of the license-plate
	Number	The license-plates are many or Nothing
	Size	Change of the size of license-plates according to the distance and type
	Font	Change of the font
	Etc	Color of the character, Background and distortion of LP
Environ-ment variations	Weathers	Rain and Snow
	Brightness	Day, night, and Lighting such as headlights
	Occlusion	It obscures the plate
	Background	Similar to those license plates

2. 기존 연구

일반적인 한국의 번호판은 직선과 색상정보를 포함하고 있는 사각형이다. 전통적인 번호판 검출 시스템들은 이러한 영상에 존재하는 번호판의 색상, 모양, 대칭, 비율과 같은 정보들을 찾고자 하였고, 이를 위하여 색상(RGB), 에지, 코너, 라인과 같은 특징 벡터들을 CCA(Connected component analysis)와 모폴로지, 에지 연산, 허프 변환 등을 이용하여 검출하였다. 그리고 이렇게 얻어진 결과들을 활용하여 번호판 검출 시스템들을 구성해왔다[1, 3-8]. 이러한 방법들은 속도의 이득이 있었으나 사전에 생각하지 못하였던 복잡한 환경에서는 사용하기 어려웠으며, 허프 변환은 상대적으로 검출률이 더 뛰어났으나, 연산량이 증가하는 문제가 있었다. 이외에도 영상 정보로부터 Haar-like, HOG와 같은 특징벡터들을 직접 생성하고 이를 다층 신경망(MLP), adaboost, SVM(Support Vector Machine)과 같은 분류기와 결합하는 연구들도 진행되어왔다[2, 10, 11, 13]. 이를 통해 기존보다 성능을 개선시킬 수 있었으나, 모델의 복잡성 증가도가 증가하여 동작 속도가 감소하며, 과적합이 발생하는 문제들이 있었다.

최근 Deep Learning을 번호판 검출에 응용하는 연구들이 늘어나고 있다[16]. CNN은 전통적인 분류기들과는 달리 기존의 특징벡터들을 활용하지 않으며, 네트워크 스스로 생성한 특징을 사용한다. 이를 번호판 검출에 활용하고, 이와 함께 경계면 추정을 위한 별도의 알고리즘을 추가함으로서 많은 성능 향상을 달성할 수 있었다. 그러나 이러한 방법은 번호판 영역의 추정을 위하여 많은 연산을 필요로 하며, 추가적인 대량의 CPU 연산으로 인하여, GPU를 활용한 효율적인 병렬 연산을 어렵게 하였다. 그렇기 때문에 더 나은 성능 향상을 위한 CNN의 깊이 조절에 영향을 주어, 검출 성능과 동작 속도 사이의 불균형을 초래하였다.

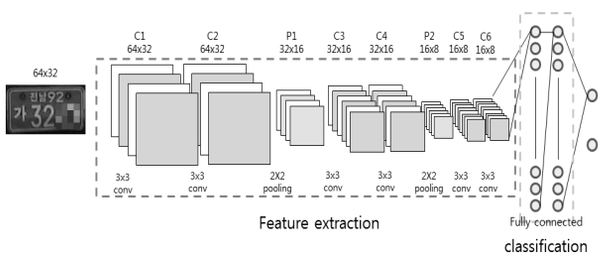

Fig. 1. The Structure of Convolutional Neural Networks (CNN) for License Plate Filtering

2.1 Convolutional Neural Networks

최근 다양한 분야에 Deep Learning이 적용되어 좋은 성과들을 보이고 있다. 이러한 Deep Learning은 Deep Belief Networks(DBN)[17], Recurrent Neural Networks(RNN) Convolutional Neural Networks(CNN)와 같이 크게 3종류 구분할 수 있다. 특히 사람의 뇌가 시각정보를 받아들이는 과정을 모방한 CNN은 영상처리 분야에서 좋은 결과를 보이고 있다[18-20]. CNN은 신경망을 기반으로 하고 있으며, Fig. 1에서와 같이 Convolution Layer와 Pooling Layer로 이루어진 '특징추출'과 Fully-Connected Layer로 이루어진 '분류'로 이루어져 있다. 기존의 물체 검출 과정은 기존의 HOG, Haar-like와 같은 전문가에 의해 설계된 특징 (Handcrafted Feature)들을 활용하여 특징 추출을 진행하였다. 이와 달리 CNN은 훈련 샘플에 대하여 최적화된 특징 벡터들을 스스로 학습하고, 이를 활용하여 특징맵(Feature Map)을 생성하여

차원을 감소시키고, 결과에 대하여 '분류'를 진행하여 모델의 복잡도 증가에 따른 과적합(overfitting)을 최소화하여 성능을 향상시키고 있다.

2.2 Faster Region Based Convolutional Neural Networks

Faster R-CNN[14]는 R-CNN[21], SPPNet[22], Fast R-CNN[23]을 확장한 구조로서 Fast R-CNN[23]과 비슷한 구조를 가지고 있으며, 이미지 전체를 입력으로 사용한다. 기존의 R-CNN 구조들은 Region Proposal을 위하여 Selective Search[24]를 사용해왔다. Faster R-CNN[14]는 이러한 Selective Search[24] 대신 Region Proposal Networks(RPN)을 제안하였다. Faster R-CNN[14]는 Fig. 2와 같이 CNN의 '특징추출'를 통해 Feature Map을 생성하고, 이를 Slide-window 방식으로 convolution 하여 벡터들을 생성한다. 이를 통해 k개의 후보에 대한 k개의 후보의 점수(2-dim: 객체일 확률, 객체가 아닐 확률)와, Bounding Box를 위한 k개의 좌표(4-dim: $< x_1, y_1 >< x_2, y_2 >$)를 생성하고, 이를 기반으로 관심영역(ROI)를 만들어 내어 연산량의 이득을 얻어낸다. 이렇게 얻어진 ROI Layer 결과에 대하여 CNN의 '분류'를 위한 Fully-Connected layer들을 적용함으로서 최종적으로 물체의 검출을 완료하게 된다.

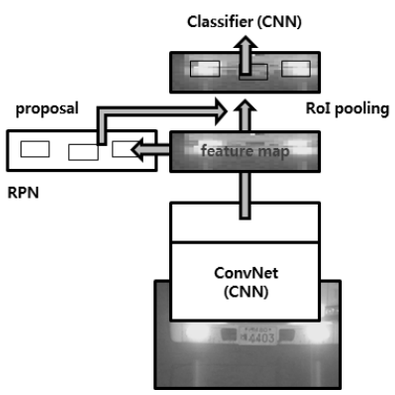

Fig. 2. The Structure of the Faster R-CNN

Region Proposal을 위하여 Selective Search[24]를 사용한 기존의 R-CNN[21], Fast R-CNN[23]들은 정확성은 높았으나, 많은 연산을 필요로 하였다. 또한 병렬 연산에 적합하지 않은 구조였기 때문에 GPU를 활용하기에 어려움이 많았다. 하지만 Faster R-CNN[14]는 이러한 Region Proposal 단계를 CNN[15, 20]의 '특징추출'과 공유함으로서, 연산량의 이득과 함께 GPU를 활용한 병렬연산의 적용을 수월하게 하였고, 이를 통해 빠르고 정확한 결과를 얻을 수 있었다.

3. 제안하는 번호판 검출 시스템

본 논문에서는 Faster R-CNN[14]과 CNN을 활용하여, 환경 변화에 강인한 실시간 자동차 번호판 검출 시스템을 제안한다. 제안하는 시스템은 별도의 전처리 과정을 없이 영상 그대로를 입력으로 활용하였다. Fig. 3은 본 논문에서 제안하는 번호판 검출 시스템의 구조도이다.

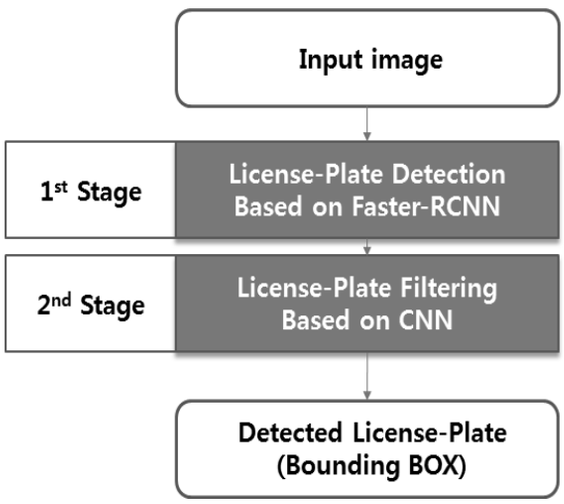

Fig. 3. The Structure of the Proposed License Plate Detection

시스템은 Faster R-CNN[14]를 활용한 번호판 '검출' 단계와 Fig. 1의 CNN을 활용한 '필터링'의 두 단계로 구성되어 있다. 첫 번째 단계인 '검출' 단에서는 영상 그대로를 Faster R-CNN[14]의 입력 영상으로 활용하여 번호판 영역의 BBOX 좌표와 번호판일 확률을 출력한다. 여기에서는 CNN[15, 20]을 내부 구조로 사용하여, 매우 높은 정확도로 TP(True Positive)를 검출할 수 있다. 하지만 훈련 데이터 및 훈련 횟수가 충분하지 못할 경우에는 TP와 함께 FP(False Positives)들이 검출된다. 본 논문에서는 이러한 FP를 효과적으로 제거하기 위하여, '검출' 단계와는 달리 충분히 많은 데이터를 사용하여 학습된 얕은 구조의 CNN을 사용하는 '필터링' 단계를 추가하였고, '검출' 단계의 결과에 대하여 번호판이 맞는지만을 판별하도록 구성하였다.

3.1 Faster R-CNN을 이용한 번호판 검출

최근 물체인식분야에서 ZFNet[15], GoogLeNet[19], VGG[20], ResNet[25] 등의 다양한 형태의 상대적으로 깊은 CNN 구조가 제안되었고, 1,000개의 물체를 분류하는 ImageNet Challenger에서 매우 높은 성능을 달성하였다. Faster R-CNN[14]는 내부 구조로 이러한 깊은 CNN[15, 19, 20, 25]들을 적용함으로써, 20개의 물체에 대하여 인식 및 위치까지 함께 추정해야하는 Pascal VOC 대회에서 빠른 동작 속도 및 검출 영역의 정교함과 함께 높은 인식률을 보여주고 있다.

본 논문에서는 전체 이미지를 입력으로 하여, 입력과 번호판의 크기에 관계없이 정확하게 번호판의 영역을 검출하고 적절한 동작속도를 가지도록 Faster R-CNN[14]을 사용하였다. 또한 내부에는 ZFNet[15]와 VGG-16[20]을 적용하였다.

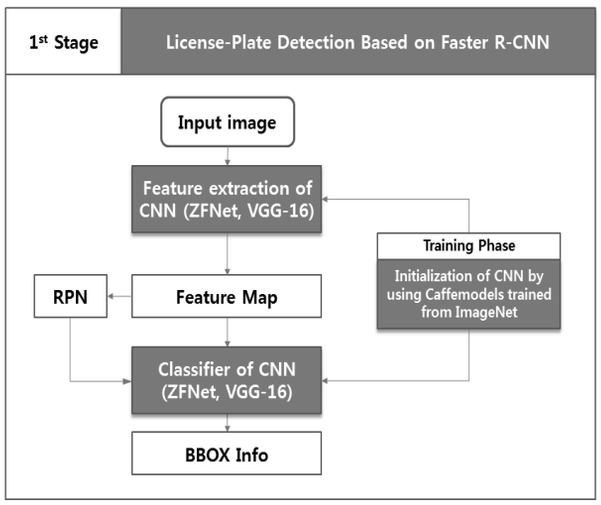

Fig. 4. The structure of License-Plate Detection Based on Faster R-CNN

3.2 CNN을 이용한 번호판 필터링

깊은 구조의 CNN[15, 20]들을 Faster R-CNN[14] 적용할 경우 정확한 자동차 번호판 영역의 검출이 가능하다. 하지만 훈련 데이터 및 훈련 횟수가 충분하지 않을 경우에는 True Positive와 함께 False Positive가 발생 하는 경우가 많이 발생한다. 본 논문에서 Faster R-CNN[14]를 위해 사용된 훈련샘플은 약 4천여장이였으며, 여기에 Negative 영상을 포함한 훈련샘플을 새롭게 추가하여 훈련시키는 것 대신에 Fig. 1과 같은 상대적으로 작은 구조의 CNN을 사용한 '필터링' 단계를 추가하였다. 여기에 사용된 CNN 구조는 상대적으로 얕은 구조의 CNN이다. 하지만 충분히 많은 Positive와 Negative 영상들을 학습시킴으로써, 효과적으로 False Positive 결과들을 제거할 수 있었다. 사용된 CNN구조는 64*32 크기를 입력으로 설정하였으며, 인수분해를 통해 연산량과 성능의 이점을 얻기 위하여, 7x7, 5x5와 같은 넓은 필터 대신 3x3의 작은 필터를 여러 번에 걸쳐 사용하도록 설계하였다[19]. 또한 Fully-Connected Layer에 Dropout [26]을 적용하여 매번 무작위로 연결을 끊음으로써 과적합을 최소화 하도록 구성하였다.

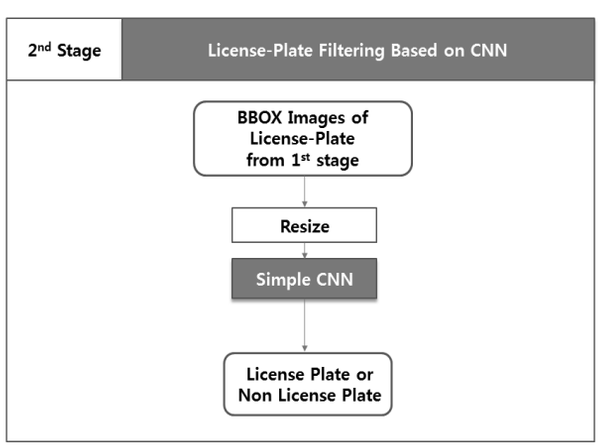

Fig. 5. The Structure of License-Plate Filtering Based on CNN

'필터링' 단계는 Fig. 5와 같이 진행되며, '검출' 결과들의 크기를 재조절하여, '필터링'의 입력으로 활용하였고, 이를 통해 번호판이 맞는지만를 확인하였다. 이에 따른 결과가 True 경우, 해당 크기를 재조절하기 전 결과를 출력하도록 하였다.

4. 실험 및 성능 평가

4.1 실험 환경

실험은 Table 2에서와 같은 환경에서 이루어졌으며, 학습 및 실험은 Linux 기반 Caffe 환경에서 Cuda 7.5와 cuDNN4 그리고 openCV 3.0을 적용하여 진행되었다.

Table 2. Experimental Environment

Processe	Intel I7-5930K@3.5GHz
Memory	64Gb
VGA	Geforce TITAN X D5 12G
OS	Ubuntu 14.04 x64
toolkit	Caffe, Cuda 7.5, cuDNN 4, OpenCV 3.0
Language	Python

4.2 데이터 셋

본 논문에서 사용된 데이터들은 Fig. 6과 같은 환경에서 두 개의 카메라를 통해 촬영된 1280*720 해상도의 이미지 80,000여장 (Table 3의 FD1-FD4) 그리고 서로 다른 시점에서 촬영된 3가지 해상도의 4,000여장의 이미지(Table 3의 FD0)들로 구성되어 있다.

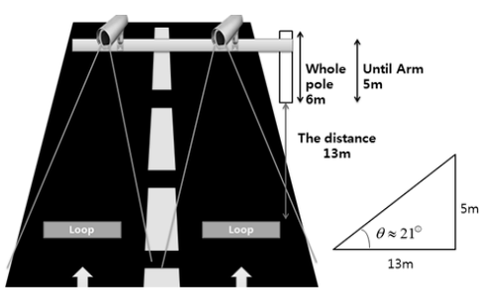

Fig. 6. Set up Environment for Data Collection

이미지들은 실제 도로를 주행 중인 자동차들을 촬영하였으며, 차량의 번호판 영역을 포함하고 있으며, 데이터셋에 대한 세부적인 정보는 Table 3과 같다.

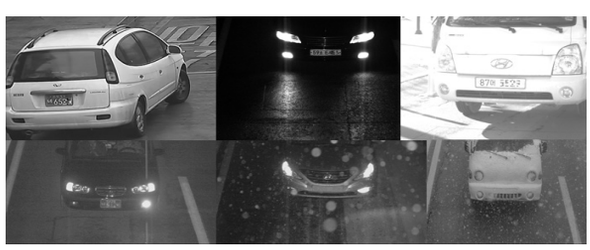

Fig. 7. Examples of License Plate Images

데이터들은 Fig. 7과 같이 시간의 흐름에 따라 촬영되었으며, 우중충한 날씨, 화창한 날씨 또는 비오는 날, 눈이 내리는 날과 같이 다양한 환경에 노출된 이미지들로서, 불균형한 조도, 크기 변화, 회전 등 Table 1과 같이 번호판 검출을 어렵게 하는 요소들을 포함하고 있다.

Table 3. Data Sets Used in the Experiments

Data sets			FD0	FD1	FD2	FD3	FD4	Total
# of images			4,212	2,000	15,991	36,086	25,661	83,950
Image feature	Reso-lution	720x480	O	X	X	X	X	
		900x600	O	X	X	X	X	
		1,280x960	X	O	O	O	O	
		1280x960	O	X	X	X	X	
	day		O	O	O	O	O	
	night		O	X	O	O	O	
	rain,snow		O	X	X	X	O	
	zoom		X	O	O	O	O	
	View point		4	1	1	2	2	

4.3 Faster R-CNN 기반 번호판 후보 검출의 훈련 및 평가

훈련 및 실험은 데이터셋의 종류에 따라서 다르게 진행되었다. 실험을 거듭할수록 전 실험에서 얻어진 잘못된 결과들의 일부를 다음 실험의 훈련샘플로 추가하였으며, 매 실험마다 새로 훈련하였다. 또한 각 평가는 TP(True Positive)만 존재할 경우를 True로 정의하였으며, TN(True Negative), FP(False Positive), FN(False Negative)가 TP와 함께 존재하는 경우들은 모두 실패한 것으로 정의하였다.

Faster R-CNN[14] 내부 구조에는 ZFNet[15], VGG-16[5]의 두 가지 종류의 CNN을 사용하여 성능을 비교하였으며, 훈련데이터의 수가 충분하지 않다고 판단되어, ImageNet의 1,240,000장에 대하여 충분히 학습되어진 CaffeModel을 초기값으로 하여 Fine-Tunning을 하였다. 훈련은 최대 30 Epoch까지 진행하였으며, 평가에는 10 Epoch, 30 Epoch까지 학습한 모델들을 사용하여 진행하였다.

1) 첫 번째 실험

첫 실험에서는 FD0의 이미지들 중에서 임의로 1,000장을 선택하여 학습하였고, 나머지 3,212를 대상으로 평가를 진행하였다. 그에 대한 결과는 Table 4와 같다. 첫 실험은 번호판에 대한 Faster R-CNN[14]의 성능을 확인하기 위하여 진행하였으며, 내부 구조는 VGG-16[20]에 대해서만 진행하였다. 이를 통해 98.38%의 높은 검출률을 확인할 수 있었다

Table 4. 1st Detection Experiment Using FD0

# of images		structure	# of Epoch	Detection Rate
training	test			
1,000	3,212	VGG-16	10	98.287%
			30	98.381%

2) 두 번째 실험

두 번째 실험은 FD0의 4,200장을 학습 데이터로 활용하여 ZFNet[15]과 VGG-16[20] 모델을 훈련시켰으며, FD1의 2,000장의 영상을 이용하여 테스트를 진행하였다. Table 5에서와 같이 1,000장에서 4,200장으로 훈련 데이터를 확장함으로써 98.38%에서 99.85%로 성능이 향상되는 것을 확인할 수 있었다. 또한 VGG-16[20]와 비교하여 상대적으로 얕은 구조의 ZFNet[15]을 사용하여도 99.70%의 높은 검출률을 얻을 수 있었다.

Table 5. 2nd Detection Experiment Using [FD0+FD1]

# of images		structure	# of Epoch	Detection Rate
training	test			
4,200	2,000	ZF	10	99.50%
			30	99.70%
		VGG-16	10	98.80%
			30	99.85%

3) 세 번째 실험

세 번째 실험에서는 먼저 FD2의 데이터를 활용하여 두 번째 실험에서 얻은 훈련 모델에 대한 테스트를 진행하였다. 이에 따른 결과는 Table 6의 첫 번째 행에서와 같이 99.50%였고, 두 번째의 99.85%와 비교하여 정확도가 0.35% 감소한 것을 확인할 수 있었다. 이때 검출에 실패한 경우들을 살펴보면 Fig. 8과 같은 사업용 대형 번호판이었으며, 이는 지금까지의 훈련 데이터에는 전혀 포함되지 않은 번호판 유형이었다.

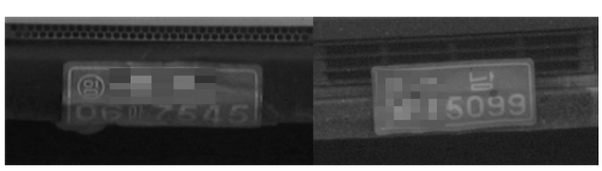

Fig. 8. Examples of Large Business License Plates

그럼에도 해당 유형의 번호판을 검출할 수는 있었으나, 일부 실패함으로써 약간의 성능하락을 가져왔다. 이러한 문제의 해결을 위하여 검출에 실패한 사업용 대형 번호판 중에서 5장을 선택하여 훈련 데이터로 추가하여 학습을 다시 하였다.

Table 6. 3rd Detection Experiment Using [Model from 2nd Experiment +FD2]

# of images		structure	# of Epoch	Detection Rate
training	test			
4200	15991	VGG-16	30	99.506%
4205 (+5)	15986	ZF	10	99.112%
			30	99.143%
		VGG-16	10	99.799%
			30	99.837%

이렇게 학습된 모델을 사용하여 FD2에 대하여 평가 결과는 Table 6의 두 번째 행과 같았다. 이를 통해 사업용 대형 번호판들을 포함한 전반적인 검출률을 개선할 수 있었다.

4) 네 번째 실험

세 번째 실험에서 Fig. 8과 같은 사업용 대형 번호판을 제외한 검출에 실패한 번호판을 살펴보면 Fig. 9와 같이 야간에 촬영된 번호판들이 많았다. FD1은 야간 촬영 영상이 전혀 포함되어 있지 않은 데이터로서 야간 영상이 포함된 FD2에 대한 검출률은 FD1보다 감소하는 것을 알 수 있었다.

Fig. 9. Example of License Plates Captured at Night

이때 실패한 야간 영상들을 살펴보면 Fig. 9 왼쪽의 한 장을 제외하고, 모든 경우에 번호판 주위에 매우 강한 조명이 존재하였다. 이러한 요인들이 번호판 검출에 영향을 주는 것으로 판단하였다. 이러한 문제의 개선을 위하여 검출에 실패한 야간 영상의 일부를 네 번째 실험의 훈련 데이터로 추가하였고, 새롭게 학습을 진행하였다. 이렇게 얻은 모델을 활용한 평가 결과는 Table 7과 같았으며, 이를 통해 강렬한 조명이 있는 번호판들에 대한 검출률을 크게 개선할 수 있었다.

Table 7. 4th Detection Experiment Using [Model from 3rd Experiment +FD3]

# of images		structure	# of Epoch	Detection Rate
training	test			
4220 (+15)	36086	ZF	10	99.639%
			30	99.869%
		VGG-16	10	99.850%
			30	99.908%

5) 다섯 번째 실험

다섯 번째 실험에서는 네 번째 실험에서 검출에 실패한 번호판 일부를 무작위로 선택하여 훈련 샘플에 추하여 새롭게 학습을 하였다. 이를 활용한 FD4에 대한 평가 결과는 Table 8과 같았다. VGG-16[20]를 사용할 경우 99.906%의 매우 높은 검출률을 얻을 수 있었다. 특히 FD4는 전체 영상 중에서 잡음이 가장 많은 데이터로서, 눈, 비, 조명 변화와 같은 요소들을 가장 많이 포함하고 있다. 그럼에도 불구하고, 이전 실험의 99.90%, 99.83%들과 비교하여 별다른 성능하락 없이 높은 검출률을 얻을 수 있었다. 또한 다섯 번째 실험에서는 Ture Positive들은 모두 검출함으로서, 이러한 모델이 환경적인 요소들에 대하여 강인함을 확인할 수 있었다.

Table 8. 5th Detection Experiment Using [Model from 4th Experiment +FD4]

# of images		structure	# of Epoch	Detection Rate
training	test			
4235 (+15)	25661	ZF	10	99.664%
			30	99.816%
		VGG-16	10	99.840%
			30	99.906%

실험 결과 대부분의 True Positive를 검출할 수 있었다. 하지만 Fig. 10과 같이 약간의 False Positive를 함께 검출하는 경우들이 있었는데, 이는 모두 잘못된 결과로 정의하였다. False Positive는 주로 광고를 위한 차량 위의 문자와 자동차의 헤드라이트 및 제조회사의 마크들이 주를 이루었다. 이러한 Fase Positive들은 실험을 거듭할수록 점차 감소하였다. 이러한 결과들을 통해 지속적인 학습데이터들의 확장을 통해 검출률을 개선할 수 있음을 확인하였다.

Fig. 10. Example of False Positives

검출되는 False Positive의 확률을 살펴보면 대부분 0.9 이하였다. 이는 임계치를 0.9 이상으로 엄격하게 설정함으로써 False Positive를 제거할 수도 있었다. 하지만 본 논문에서는 Faster R-CNN[14]의 기본 설정에 따라 임계치를 0.8로 설정하여 평가를 진행하였다. 또한 이후 최종 평가를 위하여 다섯 번째 실험에서 검출에 실패한 결과의 일부를 훈련샘플에 추가하여 4250장으로 확장하여 새롭게 훈련을 진행하였다.

4.4 CNN 기반의 번호판 필터링의 훈련 및 평가

번호판 검출과정에서 발생하는 False Positive들은 임계값의 조절과 훈련 데이터의 확장을 통해 충분히 해결 가능한 문제들이다. 하지만 본 논문에서는 임계값을 기본 설정 값에 맞췄으며, 더 이상 훈련 샘플을 증가시키지 않았기 때문에 이러한 False Positive의 제거를 위하여 번호판의 '필터링'만을 위한 Fig. 1과 같은 구조의 CNN을 뒤에 추가하였다. 여기에 사용된 데이터셋은 아래 Fig. 11과 같다.

Fig. 11. (Up) LP with Empty Space (Down) LP Without Empty Space

- 220×81 크기의 여백이 있는 번호판 : 8500장
- 다양한 크기의 여백이 없는 번호판 : 4400장
- 다양한 이미지에서 얻는 Negative : 3000장

여백이 있는 번호판은 번호판 주위 여백이 포함된 이미지로서 220x81의 고정된 크기의 이미지이며, 여백이 없는 이미지는 번호판에 맞춰 잘린 다양한 크기의 이미지들로 구성되어 있다. Negative는 배경 및 번호판과 유사한 데이터들로부터 얻어진 이미지들을 사용하였다. Negative를 제외한 이미지들은 일차적으로 64*32크기로 조정하였으며, 이렇게 얻어진 번호판과 Negative들을 Resize, Crop, Mirror를 통해 Positive : 696,600장, Negative : 1,215,500장으로 증가시켰다. 이렇게 얻어진 결과에서 20만장을 무작위로 선택하여 validation과 test에 할당하였다. 그에 대한 학습결과와 평가는 Table 9와 같다.

Table 9. Filtering Experiment

Epoch	validation	Epoch	validation
1	99.211	10	99.991
5	99.879	15	99.998
TEST		99.999	

'필터링'에 사용된 CNN의 구조는 Faster R-CNN[14]에서 사용된 ZFNet[15], VGG[20]와 비교하여 매우 얕고 작은 네트워크이다. 하지만 Faster R-CNN[14]에서와 달리 Positive와 함께 충분히 많은 Negative들을 훈련 데이터로 활용함으로써, 주어진 데이터들에 대하여 매우 정확하게 번호판을 판별할 수 있었으며, '검출' 단계에서 발생하는 False Positive들도 대부분 제거할 수 있었다.

4.5 최종 번호판 검출 평가

지금까지의 실험은 '번호판 후보영역 검출' 및 '필터링'에 대하여 서로 다른 데이터와 함께 개별적으로 진행되었다. 그렇기 때문에 더욱 정확한 결과의 확인을 위하여, 4.3과 4.4에서 얻어진 학습 모델을 사용하여, 훈련에 사용되어진 4250장을 제외한 80,000여장에 대한 '검출' 및 '필터링' 평가를 한 번에 진행하였다.

Table 10. Average Running Time in Faster R-CNN

	ZFNet	VGG-16
속도 (Avg)	68ms/image	158ms/image

또한 ZFNet[15]과 VGG-16[20]을 사용했을 때의 속도차이는 Table 10과 같이 2배 이상 차이가 나는 것을 확인할 수 있다. 이전 실험을 통해 '필터링'의 CNN을 통하여 Negative를 충분히 제거 가능하다고 판단하였고, 동작속도의 향상을 위하여 VGG-16[20] 대신 ZFNet[15]만을 Faster R-CNN[14]에 적용하여 평가를 진행하였다. '검출' 단계에서의 Faster R-CNN의 임계치는 0.8로 설정하였으며, 이를 만족하는 번호판 영역들을 '필터링'을 위한 CNN의 입력으로 활용하였다. 이후 '필터링' 단계에서 True를 만족하는 경우들에 대하여 크기조절 이전의 번호판 영역들을 최종 출력하였다. 이러한 과정들에 대한 결과는 Table 11 및 Fig. 12와 같다.

Table 11. Experimental Results of the Proposed System Using 2 Stages

# of correct detections (# of all cases)	Detection Rate
79,675/79,720	99.943%

Fig. 12. Example Results of Detecting

ZFNet[15]를 활용한 '검출' 영역의 정확도는 99.80% 였으며, 이후 '필터링'을 통해 False Positive들을 제거함으로서 99.94%의 높은 검출 결과를 얻을 수 있었다. 최종적으로 제안된 검출 시스템이 검출에 실패한 번호판들을 살펴보면 Fig. 13과 같다. 이와 같이 눈으로 덮여 있는 번호판들에 대해서 '검출' 단계에서는 번호판 영역으로 검출되었으나, 이후 '필터링' 단계에서 일부 Negative로 판단되어 제거되는 경우들이 존재하였다.

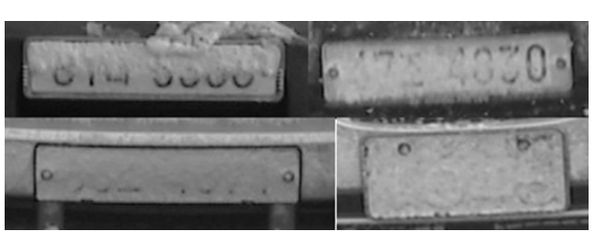

Fig. 13. Examples of Detection Failures at the Final Test

4.6 이전 시스템과의 결과 비교

Table 12는 제안된 시스템의 성능을 본 논문에서 활용한 데이터셋과 같은 데이터셋의 일부(5,200장)을 이용하여 제안되었던 Adaboost를 활용한 번호판 검출 시스템의 결과[27]

와 비교하여 보여주고 있다. 본 논문에서는 Deep Learning 기법[14, 15, 20]을 활용함으로써 True Positive 검출률을 5.65% 향상시켰으며, False Positive를 모두 제거할 수 있었다. 또한 GPU를 활용함으로써 동작속도를 50ms 이상 향상시킬 수 있었다.

Table 12. The Performance Comparison of the Proposed Method and a Previous Method Using the Same Dataset

Method	True Positive	False Positive	Avg running time
Adaboost + Segmentation filter	0.9429	0.0294	131.64 ms
Faster R-CNN + CNN	0.9994	0.0000	80 ms

이외에도 기존 시스템은 Fig. 14와 같은 왜곡들이 존재할 경우 검출률이 감소하는 문제가 있었다. 본 논문에서 이러한 유형의 번호판들을 일부 학습시킴으로서, 별다른 성능 저하 없이 번호판 영역의 검출에 성공할 수 있었다.

Fig. 14. Examples of Detection Success Under Difficult Environments

5. 결 론

본 논문에서는 Faster R-CNN[14]와 CNN을 활용하여 상대적으로 환경변화에 강인하며, 99.94%의 높은 검출률과 함께 80ms/image를 넘지 않는 실시간 번호판 검출 시스템을 제안하였다. 또한 단계적인 실험을 통해 지속적인 훈련 데이터의 확장을 통하여 쉽게 에러를 줄여나갈 수 있음을 확인하였고, 상대적으로 더 얇은 구조의 CNN을 대량의 훈련 데이터를 통해 충분히 학습함으로써 효과적으로 False Positve를 제거할 수 있음을 확인할 수 있었다. 특히 최근에는 Fig. 14의 아래 번호판과 같이 번호판 주위에 강한 조명 및 형광물질을 도포하거나, 문자영역에 형광 스티커등을 부착하는 행위들이 많아지고 있다. 이러한 문제들에 대하여 본 논문에서 제안하는 시스템은 사람의 눈으로 식별 가능한 수준의 영상 데이터에 대해서는 거의 정확하게 번호판 영역을 검출할 수 있었으며, 검출에 실패한 경우들에 대해서도 훈련 데이터의 확장을 통해 충분히 검출 가능할 것이다.

본 논문의 훈련은 8만여장의 전체 데이터와 비교하여 상대적으로 적은 4천여장의 이미지만을 Faster R-CNN[5]의 학습에 활용하였다. 적은 훈련 데이터만으로도 충분한 결과를 얻기 위하여 ImageNet을 통해 충분히 학습된 Caffe Model을 활용하였으나 이러한 Caffe Model의 활용은 Faster R-CNN [14]에 사용되는 CNN을 직접 설계하는데 장애가 되었다. 또한 훈련을 위한 Negative 데이터의 부족으로 인해, False Positive가 생성되는 문제들이 발생하였다. 이러한 문제를 해결하기 위하여 우리는 '필터링'만을 위해 상대적으로 작은 구조인 Fig. 1의 CNN을 추가하였고, 대량의 False Positive를 활용함으로써 좋은 결과를 얻을 수 있었다.

이를 통하여 이후 훈련 데이터들을 지속적으로 확장시켜, 충분히 '검출' 단계를 훈련시킨다면 '필터링'에 사용된 CNN 없이도 충분한 False Positive의 제거와 함께 Fig. 13과 같이 True Positive를 제거하는 문제를 최소화할 수 있을 것이다. 또한 충분한 데이터들을 확보함으로서 Caffemodel 대신 번호판에 최적화된 CNN구조를 제안할수 있을 것이다.

향후 연구에서는 본 논문에서 제안하는 검출 시스템을 확장하여, 환경변화에 강인하고 높은 인식률을 위한 실시간 번호판 인식 시스템을 설계하고자 한다.

References

[1] P. Viola and M. Jones "Robust Real-time Face Detection," *International Journal of Computer Vision*, Vol.57, No.2, 2004.

[2] Md. M. K. Sarker, "License Plate Detection Using Cascade Adaboost and Heuristic Energy Map," Master's degree thesis, Chonbuk National University, 2013.

[3] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.886-893, 2005.

[4] 최인수, 진문용, 박동선, "모폴로지 연산과 신명망 시스템을 이용한 차량 번호판 검출 및 숫자 인식," 스마트미디어학회, 2013 춘천 정원엑스포 ICT 합동학술대회 논문집, 제2권, 제1호, pp.58-61, 2013.

[5] M. Sarfraz, M. J. Ahmed, and S. A. Ghazi, "Saudi Arabian license plate recognition system," in *Proc. Int. Conf. Geom. Model. Graph.*, pp.36-41, 2003.

[6] A. M. Al-Ghaili, S. Mashohor, A. Ismail, and A. R. Ramli, "A new vertical edge detection algorithm and its application," in *Proc. Int. Conf. Comput. Eng. Syst.*, pp.204-209, 2008.

[7] V. Kamat and S. Ganesan, "An efficient implementation of the Hough transform for detecting vehicle license plates using DSPs," in *Proc. Real-Time Tech. Applicat. Symp.*, pp.58-59, 1995.

[8] P. Wu, H.-H. Chen, R.-J. Wu, and D.-F. Shen, "License plate extraction in low resolution video," in *Proc. Int. Conf. Pattern Recognit.*, Vol.1, pp.824-827, 2006.
[9] K. Deb and J. Kang-Hyun, "HSI color based vehicle license plate detection," in *Control, Automation and Systems, 2008. ICCAS 2008. International Conference on*, pp.687-691, 2008.
[10] L. Dlagnekov, "License Plate Detection Using AdaBoost," San Diego, CA: Computer Science and Engineering Dept., 2004.
[11] S. Z. Wang and H. J. Lee, "A cascade framework for a real-time statistical plate recognition system," *IEEE Trans. Inform. Forensics Security*, Vol.2, No.2, pp.267-282, 2007.
[12] C. D. Nguyen, M. Ardabilian, and L. Chen, "Robust Car License Plate Localization Using a Novel Texture Descriptor," *IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp.523-528, 2009.
[13] W. T. Ho, H. W. Lim, and Y. H. Tay, "Two-Stage License Plate Detection Using Gentle Adaboost and SIFT-SVM," *First Asian Conference on Intelligent Information and Database Systems*, pp.109-114, 2009.
[14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," arXiv preprint arXiv:1506.01497, 2015.
[15] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional neural networks," in *European Conference on Computer Vision (ECCV)*, 2014.
[16] Hui Li and Chunhua Shen, "Reading Car License Plates Using Deep Convolutional Neural Networks and LSTMs," arXiv arXiv:1601.05610v1 [cs.CV].
[17] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," Neural Comput., Vol.18, No. 7, pp.1527-1554, 2006.
[18] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems 25, pp.1106-1114, 2012.
[19] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. "Going deeper with convolutions," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR),*, pp.1-9, 2015.
[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *ICLR,* 2015.
[21] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region based convolutional networks for accurate object detection and segmentation," *TPAMI*, 2015.
[22] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *ECCV*, 2014.
[23] R. Girshick, "Fast R-CNN," in *IEEE International Conference on Computer Vision (ICCV)*, pp.1440-1448, 2015.
[24] J. R. Uijlings, K. E. van de Sande, T. Gevers, and A. W. Smeulders, "Selective search for object recognition," *International Journal of Computer Vision (IJCV)*, pp.154-171, 2013.
[25] He, K., Zhang, X., Ren, S., and J. Sun, "Deep residual learning for image recognition," in *CVPR*, 2016.
[26] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *JMLR*, pp.1929-1958, 2014.
[27] Moon Yong Jin, Jong Bin Park, Dongsuk Lee, Dong Sun Park, "Real-Time Vehicle License Plate Recognition System Using Adaptive Heuristic Segmentation Algorithm," in *KIPS Tr. Software and Data Eng.*, Vol.3, No.9, pp.361-368, 2014.

이 동 석
e-mail : krustmin@gmail.com
2014년 전북대학교 전자공학부(학사)
2016년 전북대학교 전자공학부(석사)
현 재 전북대학교 IT 융합연구센터 연구원
관심분야 : Deep Learning, 인공지능, 패턴인식, 영상처리

윤 숙
e-mail : syoon@mokpo.ac.kr
1993년 전북대학교 정보통신공학과(학사)
1995년 전북대학교 정보통신공학과(석사)
2006년~현 재 목포대학교 멀티미디어공학과 부교수
관심분야 : Deep Leraning, 영상처리, 멀티미디어, 패턴인식

이 재 환
e-mail : stapler75@gmail.com
2012년 전북대학교 전자정보공학부(학사)
2014년 전북대학교 전자정보공학부(석사)
2014년~현 재 전북대학교 전자정보공학부 박사과정
관심분야 : 영상처리, 패턴인식

박 동 선

e-mail : dspark@jbnu.ac.kr

1979년 고려대학교 전기전자공학과(학사)

1984년 Missouri-Columbia(석사)

1991년 Missouri-Columbia(박사)

1991년~현 재 전북대학교 전자공학부 교수

관심분야 : 신경망, 패턴인식, 영상처리, 디지털 시스템 설계