

# 합성곱 신경망 기반의 차량 번호판 인식 시스템

임성훈, 이재흥  
한밭대학교 컴퓨터공학과  
e-mail:mi131712@hanbat.ac.kr

## Convolutional Neural Network based Vehicle License Plate Recognition System

Sung-Hoon Im, Jae-Heung Lee  
Dept of Computer Engineering, Hanbat National University

### 요 약

깊은 신경망 모델을 이용한 차량 번호판 검출과 번호판 문자 인식 시스템을 제안한다. 차량 번호판 인식 시스템은 세 가지 종류의 깊은 신경망 모델로 구성된다. 기존의 영상처리 기반의 차량 번호판 검출과 문자 인식을 전부 신경망으로 대체함으로써 영상의 밝기, 회전, 왜곡 등의 변형에 강인한 성능을 얻을 수 있다. 차량 번호판 검출률은 99.3%, 문자 영역 검출률은 99%, 문자 인식률은 98.5%를 얻었다.

### 1. 서론

영상처리 기반의 차량 번호판 인식 시스템[1, 7, 9]은 번호판 검출, 문자 인식을 위한 정교한 특징 설계가 필요하다. 이러한 정교한 특징 설계에도 불구하고 회전, 밝기, 왜곡 등에 의해 번호판 검출률과 문자 인식률이 저하될 수 있다.

본 연구는 합성곱 신경망[6]을 이용하여 번호판 검출과 번호판 문자 인식을 제안한다. [6]의 방법은 성능대비 속도가 빠르며 이를 번호판 검출에 적합하도록 작게 변경하여 번호판 영역 또는 번호판 문자 영역만을 위한 모델을 제작하였다.

### 2. 관련 연구

최근에는 번호판 검출 또는 인식 문제를 특징 설계 없이 신경망을 이용하여 해결[1, 2]하려는 연구가 있다. 신경망을 이용한 방법은 경사 하강법을 통해 손실 함수를 최소화하도록 특징 추출이 된다. 따라서 다양한 환경에서 차량 데이터를 수집하였다면 회전, 밝기, 왜곡 등에 강인한 특징이 추출된 모델을 얻을 수 있다.

### 3. 차량 번호판 인식 시스템

차량 번호판 인식 시스템은 총 세 가지 종류의 모델로 구성된다.

첫 번째 모델은 영상 내의 차량 번호판 영역을 검출하며 LPD (License Plate Detection) 모델로 부른다. 두 번째 모델은 검출된 번호판 영역 내의 문자 영역을 검출하며 LPCD (License Plate Character Detection) 모델로 부른다. 마지막으로 세 번째 모델은 검출된 문자 영역들을

인식하며 LPCR (License Plate Character Recognition) 모델로 부른다.

### 3-1. LPD 모델

LPD 모델은 300×300 크기의 컬러 영상을 입력으로 사용하며 차량의 번호판 영역을 검출한다. 이전 연구[2]의 검출 모델보다 연산량을 줄였으며 검출률을 2.3% 향상시켰다. [2]에서 제안한 모델의 개량된 모델은 표 1과 같다.

<표 1> LPD 모델 정의

#	Type	Filters	Size/Stride	Output
1	Conv	32	3 × 3 / 2	150 × 150
2	Max Pool		2 × 2 / 2	75 × 75
3	Conv	64	3 × 3 / 1	75 × 75
4	Max Pool		2 × 2 / 2	37 × 37
5	Conv	64	3 × 3 / 1	37 × 37
6	Max Pool		2 × 2 / 2	18 × 18
7	Conv	128	3 × 3 / 1	18 × 18
8	Max Pool		2 × 2 / 2	9 × 9
9	Conv	256	3 × 3 / 1	9 × 9
10	Conv	128	1 × 1 / 1	9 × 9
11	Conv	256	3 × 3 / 1	9 × 9
12	Conv	128	1 × 1 / 1	9 × 9
13	Conv	256	3 × 3 / 1	9 × 9
14	Reorg 7	512	/ 2	9 × 9
15	Route 14, 13	768		9 × 9
16	Conv	256	3 × 3 / 1	9 × 9
17	Conv	30	1 × 1 / 1	9 × 9

### 3-2. LPCD 모델

LPCD 모델은 LPD 모델보다는 신경망 마지막 연산의 결과의 크기가  $30 \times 15$ 로  $9 \times 9$ 보다는 크다. LPD 모델의 경우 차량과 배경이 함께 촬영된 영상에서의 검출을 고려하였기 때문에 차량의 번호판 영역은 영상 전체에 비해서 상대적으로 작다. 하지만 LPCD 모델은 차량의 번호판 영역이 입력으로 사용되고 번호판 문자가 번호판 영역에서 차지하는 비율이 크기 때문에 실험적으로  $30 \times 15$ 의 크기를 갖도록 구성하였다.

<표 2> LPCD 모델 정의

#	Type	Filters	Size/Stride	Output
1	Conv	32	$3 \times 3 / 1$	$240 \times 120$
2	Max Pool		$2 \times 2 / 2$	$120 \times 60$
3	Conv	64	$3 \times 3 / 1$	$120 \times 60$
4	Max Pool		$2 \times 2 / 2$	$60 \times 30$
5	Conv	128	$3 \times 3 / 1$	$60 \times 30$
6	Conv	64	$1 \times 1 / 1$	$60 \times 30$
7	Conv	128	$3 \times 3 / 1$	$60 \times 30$
8	Max Pool		$2 \times 2 / 2$	$30 \times 15$
9	Conv	256	$3 \times 3 / 1$	$30 \times 15$
10	Conv	128	$1 \times 1 / 1$	$30 \times 15$
11	Conv	256	$3 \times 3 / 1$	$30 \times 15$
12	Reorg 7	512	$/ 2$	$30 \times 15$
13	Route 12, 11	768		$30 \times 15$
14	Conv	256	$3 \times 3 / 1$	$30 \times 15$
15	Conv	30	$1 \times 1 / 1$	$30 \times 15$

### 3-3. LPCR 모델

LPCR 모델은 총 95개(가~마, 거~머, 고~모, 구~무, 버~피, 보~포, 부~후, 그~호, 바~하, 허, 호, 0~9, 울, 산, 대, 인, 천, 광, 경, 진, 기, 강, 원, 충, 북, 남, 제)의 문자를 분류한다.

실험 결과 문자의 가로, 세로의 크기가 최소  $64 \times 64$  이상을 유지하는 것이 인식률 향상에 도움이 되었다. 따라서 문자 인식 모델의 입력은  $64 \times 64$ 의 크기를 갖는 흑백 문자 영상을 사용한다.

<표 3> LPCR 모델 정의

#	Type	Filters	Size/Stride	Output
1	Conv	32	$3 \times 3 / 1$	$64 \times 64$
2	Max Pool		$2 \times 2 / 2$	$32 \times 32$
3	Conv	64	$3 \times 3 / 1$	$32 \times 32$
4	Max Pool		$2 \times 2 / 2$	$16 \times 16$
5	Conv	128	$3 \times 3 / 1$	$16 \times 16$
6	Max Pool		$2 \times 2 / 2$	$8 \times 8$
7	Conv	128	$3 \times 3 / 1$	$8 \times 8$
8	Conv	64	$1 \times 1 / 1$	$8 \times 8$
9	Conv	128	$3 \times 3 / 1$	$8 \times 8$
10	Max Pool		$2 \times 2 / 2$	$4 \times 4$
11	FC			512
12	FC			95

## 4. 차량 번호판 데이터

이번 장에서는 LPD, LPCD, LPCR 모델들을 학습하기 위한 데이터에 대해 설명한다. 각 모델들은 학습시 서로 연관되어 있지 않으며 독립적으로 학습한다.

### 4-1. LPD 학습 데이터

번호판 영역 검출 데이터는 그림 1과 같이 밝기, 위치, 날씨, 차량 종류, 배경 등 다양한 조건이 포함되도록 구성하였으며 대략 3,000장으로 구성하였다.

모든 데이터의 가로와 세로의 비율은 1:1이며 번호판 영역의 위치를 좌측 상단 좌표와 우측 하단 좌표로 저장하였다.

LPD 모델을 학습할 때에는 잡음, 밝기, 색상 변화 등의 무작위 변형을 통해 대략 15,000장으로 늘린 후 학습을 진행하였다.



(그림 1) 번호판 영역 검출 데이터

### 4-2. LPCD 학습 데이터

번호판 문자 영역 검출 데이터는 그림 2와 같이 대략 2,000장을 수집하였다. 데이터 내에는 약간의 기울기, 밝기, 다양한 종류의 번호판이 포함되어있다.

모든 데이터의 가로와 세로 비율은 2:1이며 번호판 내에 있는 모든 문자의 위치를 좌측 상단 좌표와 우측 하단 좌표로 저장하였다.

LPCD 모델을 학습할 때에는 잡음, 밝기, 색상 변화 등의 무작위 변형을 통해 대략 10,000장으로 늘린 후 학습을 진행하였다.



(그림 2) 번호판 문자 영역 검출 데이터

4-3. LPCR 학습 데이터

본 연구에서는 16종류의 차량 번호판 문자 시트를 제작하였으며 표 4와 같이 다양한 종류의 글꼴로 구성되어 있다. 한컴오피스 2014 VP의 한글 프로그램<sup>1)</sup>에 내장된 글꼴을 사용하였다.

<표 4> 차량 번호판 문자 시트

#	문자	숫자
1	궁서체(26, 굵음)	Ahn Bold(40)
2	문체부 훈민정음체 (26, 굵음)	Bahnschrift Light(40)
3	Ahn Bold(32)	Bahnschrift Light(40)
4	#Senaru(26, 굵음)	Cucumber(40, 굵음)
5	Corn(26, 굵음)	Konghan Bold (32, 굵음)
6	CouierNew(26, 굵음)	Penheullim(32, 굵음)
7	Cucumber(26, 굵음)	Cucumber(26, 굵음)
8	YJ Bonmokgak M (26, 굵음)	YJ Bonmokgak M (26, 굵음)
9	Ganeunhan(32, 굵음)	Harlow Solid Italic(32)
10	문체부 제목 돌음체(26, 굵음)	Konghan Bold (32, 굵음)
11	Headline(26)	Kidney bean(26, 굵음)
12	Human Yet(26, 굵음)	Kidney bean(26, 굵음)
13	Konghan Bold (32, 굵음)	Konghan Bold (32, 굵음)
14	Kristen ITC(26, 굵음)	Kidney bean(26, 굵음)
15	Peach(30, 굵음)	Kidney bean(26, 굵음)
16	Sinmyeong Dinaru (26, 굵음)	Kidney bean(26, 굵음)

그림 3은 표 4에서 10번 시트에 해당한다. 각 시트마다 문자의 크기와 숫자의 크기, 문자의 글꼴, 숫자의 글꼴이 다르며 글꼴의 굵음 처리 여부가 표시되어 있다.

실제 차량 번호판의 글꼴은 굵기 때문에 굵은 글꼴을 위주로 선택하였으며 번호판의 글자 형태와 비슷한 글꼴을 선택하였다.



(그림 3) 번호판 문자 글꼴 예시

차량 번호판 문자의 학습 데이터를 생성할 때에는 16종류 시트의 격자 경계 안에 있는 모든 문자들을 잘라내어서 64×64 크기로 생성하였다.



(그림 4) 문자 무작위 변형 예시

문자 인식 모델의 일반화 성능을 높이기 위해서 시트로부터 잘라낸 문자를 무작위(크기 변환, 평행 이동, 잡음, 회전, 필터 처리)로 변형하였으며 대략 300,000장의 데이터를 구성하였다.

그림 4에서 상단은 검은색 글씨와 흰색 배경이고 하단은 흰색 글씨와 검은색 배경이다. 두 가지 종류의 데이터를 생성한 이유는 차량 번호판의 종류가 흰색, 노란색, 초록색이 존재하기 때문이다. 이를 이진화[10]하면 흰색 배경의 번호판은 검은색 글씨와 흰색 배경을 갖는 이진화 이미지를 얻지만 다른 경우는 그렇지 않다. 두 종류의 이진화 이미지가 입력으로 사용되어도 인식할 수 있도록 두 가지 종류의 데이터를 생성하였다.

1) <https://www.hancom.com>

## 5. 학습 방법

본 논문에서 제안한 모든 모델의 Conv, FC 타입 연산은 Leaky ReLU ( $\alpha=0.1$ ) 활성화 함수[3]와 Batch Normalization[11]을 적용하였고 모델의 마지막 연산은 활성화 함수를 적용하지 않았다.

### 5-1. LPD 학습 방법

경사 하강법을 적용하였고 학습률은 0.0005, weight decay는 0.0005, momentum은 0.9를 적용하였다.

학습 데이터의 mini-batch 크기는 64개씩 랜덤으로 적용하였으며 학습률은 100, 500, 6,000, 8,000회 마다 각각 10, 0.2, 0.1, 0.1배로 변경하였고 총 10,000회 학습하였다.

### 5-2. LPCD 학습 방법

경사 하강법을 적용하였고 학습률은 0.0001, weight decay는 0.0005, momentum은 0.9를 적용하였다.

학습 데이터의 mini-batch 크기는 64개씩 랜덤으로 적용하였으며 학습률은 100, 500회 마다 각각 10, 0.1배로 변경하였고 총 3,000회 학습하였다.

### 5-3. LPCR 학습 방법

경사 하강법을 적용하였고 학습률은 0.1, weight decay는 0.0005, momentum은 0.9를 적용하였다.

학습 데이터의 mini-batch 크기는 64개씩 랜덤으로 적용하였으며 학습률은 1,000, 8,000, 15,000회 마다 각각 0.1, 0.1, 0.1배로 변경하였고 총 20,000회 학습하였다.

## 6. 결론

차량 번호판 인식 문제를 오로지 신경망 학습으로 해결하기 위해 세 가지 신경망 모델을 제안하였다. LPD, LPCD, LPCR 신경망 모델들을 통해 각각 번호판 영역 검출, 번호판 문자 영역 검출, 번호판 문자 영역 인식 총 세 단계를 거쳐서 차량의 번호판을 인식한다.

제안한 방법을 통해 차량 번호판 검출률은 99.3%, 문자 영역 검출률은 99%, 문자 인식률은 98.5%를 얻었으며 최종 처리 속도는 NVIDIA GTX1070 GPU에서 대략 39ms이다.

## 7. 향후 연구

본 논문에서 제안한 방법은 차량 번호판 학습 데이터와는 전혀 다른 환경에서도 검출률과 인식률이 실험 데이터의 수치와 비슷하였다. 하지만 모바일 GPU가 탑재된 소형 임베디드 보드에서 실시간 작동이 제한적이다.

향후 연구는 LPD 모델과 LPCD 모델을 [6]의 방식을 사용하는 것이 아닌 Cascaded 검출 방식[4]과 연산 비트수를 줄이는 Quantization[5, 8] 신경망을 구성하여 고성능 임베디드 플랫폼에서 실시간으로 작동될 수 있도록 연구할 예정이다.

## 감사의 글

이 논문은 2015, 2016년도 정부(교육부, 미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 지역혁신타인력양성사업(No. 2015H1C1A1035818)과 지역신산업선도인력양성사업(No. 2016H1D5A1911149)의 성과임

## 참고문헌

- [1] 임성훈, 박병주, 이재홍, “심층 컨볼루션 신경망을 이용한 번호판 인식 시스템”, 한국정보처리학회 춘계학술발표대회, 23권 제 1호, 2016.04.
- [2] 임성훈, 박시홍, 이재홍, “완전 컨볼루션 신경망 기반의 차량 번호판 검출기”, 한국정보처리학회 춘계학술발표대회, 24권 제 1호, 2017.04.
- [3] B. Xu, N. Wang, “Empirical Evaluation of Rectified Activations in Convolution Network”, arXiv preprint arXiv:1505.00853, 2015.
- [4] D. Chen, S. Ren, Y. Wei, X. Cao, and J. Sun, “Joint cascade face detection and alignment”, In Proc. ECCV, 2014.
- [5] J Chenzhuo Zhu, Song Han, Huizi Mao, and William J Dally, “Trained ternary quantization”, arXiv preprint arXiv:1612.01064, 2016.
- [6] J. Redmon, A. Farhadi, “YOLO9000: Better, Faster, Stronger”, arXiv preprint arXiv:1612.08242, 2016. 1
- [7] M. Ko, “Effective license plate character recognition based on geometric invariant features”, Ph. D. dissertation, Kyungpook National University, 2004.
- [8] M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi, “Xnornet:Imagenet classification using binary convolutional neural networks”, arXiv preprint arXiv:1603.05279, 2016.
- [9] M. Yu and Yong Deak Kim, “An approach to Korean license plate recognition based on vertical edge matching”, Proceedings of the IEEE, Vol.4, pp.2975-2980.
- [10] N. Otsu, “A threshold selection method from gray level histograms”, IEEE Trans. Syst. Man Cybern. SMC-9, 62-66, 1979.
- [11] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift”, arXiv preprint arXiv:1502.03167, 2015.