

# 준지도학습을 이용한 차량 번호판 인식 모델의 설계와 구현

김형중\*, 김정태\*, 김소연\*\*\*\*, 김수인\*\*, 민성현\*\*\*, 조성환\*\*\*, 전영진\*\*\*\* 김원겸\*\*\*\*  
\*단국대학교 소프트웨어학과, \*\*단국대학교 인공지능 융합학과,  
\*\*\*주케이사인, \*\*\*\*(주)에이아이딥

## License Plate Detection and Recognition Model based on Semi-supervised Learning

Hyungjung Kim\*, Kyungtae Kim\*, Sooin Kim\*\*, Seonghyun Min\*\*\*, Seonghwan Cho\*\*\*,  
Youngjin Jeon\*\*\*\*, Soyeon Kim\*\*\*\*, Wongyum Kim\*\*\*\*  
Department of Software, Dankook University\*,  
Department of Artificial Intelligence Convergence, Dankook University\*\*,  
KSign Inc.\*\*\*, AiDEEP Co.,Ltd\*\*\*\*  
E-mail : \*gudwnd1113@gmail.com, \*nonparanoid1@gmail.com

### Abstract

Practical deep learning technology is being used to detect and recognize license plates for intelligent vehicle control. OCR models have been based on supervised-learning methods, but labeling tasks for collected data consumes a lot of cost. This paper proposes a license plate detection and recognition model through a semi-supervised learning approach. The proposed model uses a feature extraction layer based on VGG-16 for both detection and recognition and uses not only labeled data but also unlabeled data. In the experiment, the proposed approach outperforms the comparative models in terms of data size and generalization capability. The proposed strategy can be extendible to various tasks such as video stream analysis, or potential labeler for a large-scale data.

### I. 서론

자동차 번호판의 탐지 및 인식은 지능형 차량 관제 시스템에서 중요한 연구분야로 주차 제어, 교통 관리 등 다양한 분야에서 활용된다[1]. 영상으로부터 번호판을 인식 시, 왜곡된 번호판으로 인해 세그멘테이션 및 인식의 정확도가 하락한다. 이를 해결하기 위해 딥러닝 알고리즘을 활용한 연구가 도입되고 있다[2].

딥러닝 기반 번호판 인식 기술은 객체 탐지와 문자

인식을 수행하는 2 단계 방식으로 진행된다. 객체 인식과 문자 인식 모델은 학습 데이터의 레이블 정보를 이용해 신경망의 추론과 실제 정답 간 오차가 작아지도록 학습시키는 지도학습 방식으로 훈련한다. 지도학습 방식의 훈련은 학습 데이터셋 준비 과정에서 데이터 수집과 더불어 수집된 만큼의 추가적인 레이블링 과정이 수반된다.

본 논문에서는 수집된 이미지 보다 적은 양의 레이블 정보로 번호판 탐지 및 인식을 진행하는 Pseudo label[3] 기반의 준지도학습 모델을 설계하고 평가한다. VGG-16[4]에 기반한 탐지기와 인식기로 구성된 제안 모델은 레이블 정보가 있는 데이터와 레이블 정보가 없는 이미지를 동시에 활용하여 학습할 수 있다. 제안된 준지도학습 기반의 번호판 인식 모델은 동일한 크기의 레이블 정보를 가진 학습 데이터에 대해 지도학습 방식을 이용한 모델과 비교하여 개선된 성능을 보인다.

### II. 관련 연구

기존에 제안된 번호판 인식 모델의 자동차 번호판 탐지는 기계학습, 딥러닝, 영상처리 등 기법 등이 적용되었다. 탐지된 객체로부터의 문자 또는 숫자에 대한

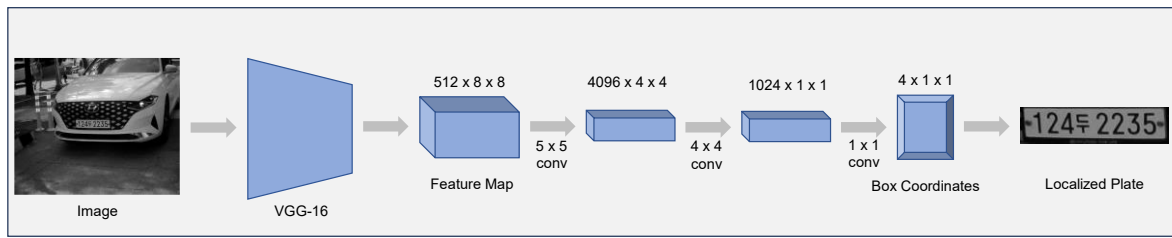


그림 1 VGG-16 기반 번호판 탐지 모델의 흐름도

인식 모델은 딥러닝 모델이 선택되었다. SVM(support vector machine, SVM)[5]을 적용하여 번호판을 탐지하고 탐지된 모델에 대하여 CNN(convolutional neural network, CNN)[6] 알고리즘을 사용하는 모델이 제안되었다[7]. 제안 모델은 Canny 윤곽선 검출 알고리즘을 이용하여 번호판 후보 이미지들을 탐지하고 사전에 지도학습된 SVM을 적용하여 번호판 영역을 탐지한다. 번호판 객체의 가로세로 영역 히스토그램을 계산하여 문자와 숫자들을 분할한다. 종횡비 기반이 아닌 탐지 위치를 기준으로 인식하여 번호판 규격이 바뀌어도 탐지 가능하였다. 분할된 문자와 숫자는 CNN 모델을 훈련하여 인식하였다.

탐지 단계에서 YOLOv4[8]를, 인식 단계에서 CNN을 이용한 모델이 제안되었다. 제안 모델은 번호판 탐지, 번호판 영상 처리, 번호판 인식의 단계를 수행한다[9]. 자음과 모음이 분리되어 있어 오인식이 잦은 한글의 인식률을 향상시키기 위해 영상 침식, 분리된 자음과 모음 영역 탐지, 음절 영역 추출 등 3 가지 영상 처리 알고리즘을 사용한다.

오픈소스 라이브러리 easyOCR은 탐지 단계에서 CRAFT(character region awareness for text detection, CRAFT) 모델을 사용한다[10]. 글자의 지역 임계값과 글자 간 유연 임계값을 설정하여 문자를 단어 단위 경계 박스로 구분한다. 인식 단계에서는 BiLSTM(bidirectional long short-term memory, BiLSTM)[11]을 사용하며 경계 박스 내 문자 수와 무관하게 인식한다.

### III. 제안 방법

#### 3.1 번호판 객체 탐지

256 × 256크기의 단일채널 흑백 이미지에 존재하는 차량에 대하여 번호판 탐지를 위한 지역화(localization)를 수행한다. VGG-16 으로부터 이미지의 특징맵을 추출

하고 합성곱 연산을 수행하여 자동차 번호판 영역을 추출한다.

그림 1은 입력 영상으로부터 번호판 객체를 추출하는 과정이다. 입력 영상에 대한 VGG-16의 출력은 512 × 8 × 8의 크기이며 채널 출력층의 완전연결계층은 512 채널의 특징맵을 경계 박스 좌표를 나타내는 4 채널의 벡터로 축소한다. 계산 복잡도를 줄이기 위해 합성곱 연산을 3 번에 나누어 거쳐 차원 축소를 수행한다. 모델의 출력층에서는 4 개 채널의 1 × 1 벡터가 출력되며 각 벡터는 경계 박스의 좌상단 x 좌표, 좌상단 y 좌표, 너비, 그리고 높이이다. 해당 좌표 정보로부터 번호판 영역을 추출한다.

#### 3.2 문자 세그멘테이션

세그멘테이션을 통해 이미지를 픽셀 단위로 분할하여 각 픽셀을 서로 다른 객체 또는 배경으로 할당한다. 번호판 내의 문자들을 인식하기 위해 번호판 이미지 내의 각 문자들을 단일 이미지로 분리한다. 이미지 이진화, 모폴로지 연산, 연결 요소 분석의 영상처리 기법을 적용하여 객체의 영역을 분리한다.

이진화는 각 픽셀을 0 과 1 두가지 값만으로 변환한다. 픽셀값을 바꾸는 것은 임계값에 의해 결정되며 제안방법에서는 이미지 내 픽셀값의 평균을 임계값으로 설정하여 이진화를 진행한다. 번호판은 이진화를 통해 문자 영역이 1의 픽셀값을, 그 외 영역이 0의 픽셀값을 가지는 단일채널 이미지가 된다.

이진화된 이미지는 얼룩이나 원근에 의한 왜곡으로 인해 둘 이상의 문자가 겹쳐 나타날 수 있다. 픽셀이 연결된 문자들을 분리하기 위해 침식(erosion)과 팽창(dilation) 두가지 모폴로지 연산을 수행한다. 2 × 2커널을 이용한 모폴로지 연산을 통해 겹쳐진 문자들을 분리시킨다.

한글은 숫자와 달리 형태소 단위로 객체가 분리되

어 하나의 문자로 탐지되지 않을 수 있다. 연결요소 분석을 통해 탐지된 객체들로부터 각 객체의 높이, 너비, 대각 길이를 고려하여 하나의 문자로 구성시킨다.

### 3.3 문자 인식

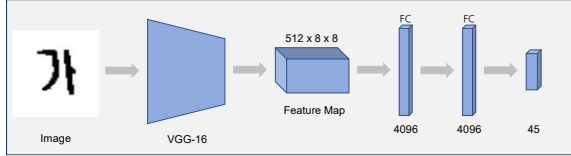


그림 2 VGG-16 기반 문자 인식 모델

그림 2 는 문자 인식 모델 구조이다. VGG-16 기반 특징 추출 계층을 통해  $64 \times 64$ 크기의 입력 데이터로부터 특징맵을 추출하고 문자를 분류한다. 번호판의 문자는 0 부터 9까지의 숫자 10자와 “가, 나, 다, 라, 마, 거, 너, 더, 러, 머, 버, 서, 어, 저, 고, 노, 도, 로, 모, 보, 소, 오, 조, 구, 누, 두, 루, 무, 부, 수, 우, 주, 하, 허, 호”로 구성된 한글 35 자로 구성된다. 모델의 출력층은 3 개의 완전 연결 계층이며 차량 번호판에 사용되는 45 개의(문자 숫자 갯수 추가) 문자를 분류하기 위해 마지막 출력 계층은 45 개의 뉴런으로 구성된다.

### 3.4 Pseudo label 기법

제안 모델은 문자 인식 단계에서 Pseudo label 기법을 활용한다. Pseudo label 은 지도학습 기반의 분류 모델 학습에 있어 레이블 정보가 없는 데이터를 같이 활용하는 준지도학습 기법이다[3]. 학습 과정 중 레이블 정보가 없는 데이터의 클래스를 분류하고 예측된 클래스를 다음 학습 단계 때 레이블 정보가 있는 데이터의 레이블  $y'_i$  로서 활용한다. 레이블 정보가 없는 데이터에 대해서 최대 예측 확률값  $f_{i'}(x)$  을 가진 클래스를 레이블로 매핑한다.

$$y'_i = \begin{cases} 1, & \text{if } i = \operatorname{argmax}_{i'}, f_{i'}(x) \\ 0, & \text{otherwise} \end{cases} \quad \text{식(1)}$$

Pseudo label 을 통해 학습 시, 학습 초기에는 레이블 정보를 가진 데이터만을 사용해 지도학습을 수행한다. 레이블 정보를 가진 데이터로 일정 에폭 동안 사전학습을 진행한 다음 레이블 정보가 없는 데이터를 추가로 활용하여 준지도학습을 수행한다.

$$L = \frac{1}{n} \sum_{m=1}^n \sum_{i=1}^C L(y_i^m, f_i^m) + \frac{1}{n'} \sum_{m=1}^{n'} \sum_{i=1}^C L(y_i'^m, f_i'^m) \quad \text{식(2)}$$

분류 문제에서의 손실함수로는 교차 엔트로피 오차 (cross entropy error)를 사용한다. 준지도학습 단계에서의 손실함수는 레이블이 있는 데이터의 교차 엔트로피 오차  $L(y_i^m, f_i^m)$ 와 레이블이 없는 데이터의 교차 엔트로피 오차  $L(y_i'^m, f_i'^m)$ 의 평균값이다. 레이블 정보가 없는 문자 이미지의 모델 출력은 학습이 진행됨에 따라 하나의 문자를 나타내는 레이블로 수렴하는 원-핫 벡터(one-hot vector)로 표현된다. 레이블 정보가 없는 데이터들에 대한 결정 경계(decision boundary)는 데이터들의 확률밀도가 적은 곳에 위치하며 데이터들이 잘 구분될 수 있다.

## IV. 실험 결과

데이터셋은 학습 데이터셋 2,000 장과 평가 데이터셋 3,746 장으로 구성하며 모델 평가지표는 정확도이다. 탐지 모델의 출력인 경계 박스 내에서 탐지된 객체의 개수가 6 개 이상일 시 정탐으로 인식한다. 번호판에 사용되는 문자 개수가 6 개 혹은 7 개이며 추가적으로 객체로 탐지되는 잡음 등을 고려하여 영역 내 숫자와 문자가 6 개 이상인 경우이다. 표 1 은 평가 데이터셋에 대한 탐지 성능 실험 결과이다. 평가 데이터셋에 대해 108 건의 오탐을 제외한 97.11%의 정확도를 보였다. 학습 데이터셋에서 신형 번호판과 구형 번호판 비율의 불균형이 존재해 상대적으로 학습 데이터가 적은 구형 번호판에서 대부분의 오탐이 발생하였다.

표 1 탐지 모델 정확도 평가

Phase	Data	Error	Accuracy (%)
Localization	3,746	108	97.11

제안된 문자 인식 모델에서의 지도학습과 준지도학습 간 성능 비교를 위해 레이블 정보를 가진 데이터의 수에 따라 세 가지 경우로 나누어 실험을 진행하였다. 학습 데이터셋에서, 각 실험에서 사용된 레이블 정보를 가진 데이터의 수는 각각 1,575 장, 2,925 장, 3,600 장이며 레이블 정보가 없는 데이터는 12,360 장으로 모든 실험에서 동일하다. 평가 데이터셋은 자동차 번호판 사진 14,471 장에서 추출한 문자 이미지로 구성하였다.

지도학습의 경우 레이블 데이터만을 사용하여 100 번 학습하였다. 반면, 준지도학습의 경우, 레이블 데이터만으로 30 번 학습을 한 뒤, 레이블 데이터와 레이블이 없는 데이터를 동시에 70 번 학습을 하였다. 표 2 는 지도학습과 준지도학습 실험 결과를 비교한 것이다. 동일한 크기의 레이블 정보를 가진 데이터셋에 대해 준지도학습 방식이 지도학습 방식보다 높은 정확도를 보인다.

표 2 학습 데이터 수에 따른 인식 모델 정확도

Labeled Data Method	1,575	2,925	3,600
Supervised	91.84%	94.22%	96.08%
Semi-Supervised	95.78%	99.89%	99.94%

표 3 은 동일한 시험 데이터셋에 대해 제안 모델과 기존 지도학습 기반 문자인식 모델의 정확도를 비교한 것이다. 비교 모델들은 사전학습된 가중치를 이용하였으며 제안 모델이 비교 모델에 비해 17 ~ 40%개선된 성능을 보인다.

표 3 자동차 번호판 인식 모델 정확도 비교

Method	Accuracy (%)
easyOCR	78.02
paddle	82.9
Pororo	73.46
SSD	66.5
SVM	53.52
Ours	99.94

## V. 결론

본 논문에서 레이블 정보가 없는 데이터를 분류 모델 학습에 사용하는 준지도학습 기반의 번호판 탐지 및 인식 모델을 제안하였다. 제안 모델은 지도학습 기반 모델과 비교하여 향상된 성능을 보였으며 이러한 준지도 학습 방식은 딥러닝 모델 구축에 있어 고비용이 소모되는 학습 데이터 수집과 레이블링 작업에 대한 소요를 줄일 수 있다. 제안 모델은 번호판뿐만 아닌 비디오 스트림, 주민등록증 등 문자 인식이 필요한 기타 영역에 활용할 수 있다. 또한 제안모델은 준지도학습 방법을 통해 레이블 정보가 없는 데이터셋에 레이블을 부여

하는 레이블러(labeler)로서 이용될 수 있다.

## 참고문헌

- [1] Jin, Moon Yong, et al. "Real-Time Vehicle License Plate Recognition System Using Adaptive Heuristic Segmentation Algorithm." *KIPS Tr. Software and Data Eng* 3.9, 2014, 361-368.
- [2] Li, Hui, and Chunhua Shen. "Reading car license plates using deep convolutional neural networks and LSTMs." *arXiv preprint arXiv:1601.05610*, 2016.
- [3] Lee, Dong-Hyun. "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks." *Workshop on challenges in representation learning, ICML*. Vol. 3. No. 2. 2013.
- [4] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556*, 2014.
- [5] Hearst, Marti A., et al. "Support vector machines." *IEEE Intelligent Systems and their applications* 13.4, 1998, 18-28.
- [6] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems* 25, 2012.
- [7] Kim, Jung-Hwan, and Joonhong Lim. "License Plate Detection and Recognition Algorithm using Deep Learning." *Journal of IKEEE* 23.2, 2019, 642-651.
- [8] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." *arXiv preprint arXiv:2004.10934*, 2020.
- [9] Kim, Jae-Jung, and Chang-Bok Kim. "Implementation of Robust License Plate Recognition System using YOLO and CNN." *Journal of Korean Institute of Information Technology* 19.4, 2021, 1-9.
- [10] Baek, Youngmin, et al. "Character region awareness for text detection." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
- [11] Huang, Zhiheng, Wei Xu, and Kai Yu. "Bidirectional LSTM-CRF models for sequence tagging." *arXiv preprint arXiv:1508.01991*, 2015.