Journal of the Korea Convergence Society
Vol. 12. No. 1, pp. 29-35, 2021

ISSN 2233-4890 / eISSN 2713-6353
https://doi.org/10.15207/JKCS.2021.12.1.029

딥러닝과 의미론적 영상분할을 이용한 자동차 번호판의 숫자 및 문자영역 검출

이정환
안동대학교 전자공학과 교수

Detection of Number and Character Area of License Plate Using Deep Learning and Semantic Image Segmentation

Jeong-Hwan Lee
Professor, Department of Electronic Engineering, Andong University

요 약 자동차 번호판 인식은 지능형 교통시스템에서 핵심적인 역할을 담당한다. 따라서 효율적으로 자동차 번호판의 숫자 및 문자영역을 검출하는 것은 매우 중요한 과정이다. 본 연구에서는 딥러닝과 의미론적 영상분할 알고리즘을 적용하여 효과적으로 자동차 번호판의 번호영역을 검출하는 방법을 제안한다. 제안된 방법은 화소 투영과 같은 전처리과정 없이 번호판 영상에서 바로 숫자 및 문자영역을 검출하는 알고리즘이다. 번호판 영상은 도로 위에 설치된 고정 카메라로부터 획득한 영상으로 날씨 및 조명변화 등을 모두 포함한 다양한 실제 상황에서 촬영된 것을 사용하였다. 입력 영상은 색상변화를 줄이기 위해 정규화하고 실험에 사용된 딥러닝 신경망 모델은 Vgg16, Vgg19, ResNet18 및 ResNet50이다. 제안방법의 성능을 검토하기 위해 번호판 영상 500장으로 실험하였다. 학습을 위해 300장을 할당하였으며 테스트용으로 200장을 사용하였다. 컴퓨터모의 실험결과 ResNet50을 사용할 때 가장 우수하였으며 95.77% 정확도를 얻었다.

주제어 : 딥러닝, 합성곱 신경망(CNN), 의미론적 분할, 자동차 번호판, 영상분할 및 인식

Abstract License plate recognition plays a key role in intelligent transportation systems. Therefore, it is a very important process to efficiently detect the number and character areas. In this paper, we propose a method to effectively detect license plate number area by applying deep learning and semantic image segmentation algorithm. The proposed method is an algorithm that detects number and text areas directly from the license plate without preprocessing such as pixel projection. The license plate image was acquired from a fixed camera installed on the road, and was used in various real situations taking into account both weather and lighting changes. The input images was normalized to reduce the color change, and the deep learning neural networks used in the experiment were Vgg16, Vgg19, ResNet18, and ResNet50. To examine the performance of the proposed method, we experimented with 500 license plate images. 300 sheets were used for learning and 200 sheets were used for testing. As a result of computer simulation, it was the best when using ResNet50, and 95.77% accuracy was obtained.

Key Words : Deep Learning, Convolution Neural Network(CNN), Semantic Segmentation, License Plate, Image Segmentation and Recognition

*This work was supported by a Research Grant of Andong National University.
*Corresponding Author : Jeong-Hwan Lee(jhlee@anu.ac.kr)
Received November 9, 2020 Revised December 10, 2020
Accepted January 20, 2021 Published January 28, 2021

1. 서론

최근 효율적인 교통신호 제어를 위해 지능형 교통제어 시스템에 대한 연구가 활발하게 진행되고 있다[1]. 특히 차량의 출입통제 및 적절한 관리, 범죄차량의 검색, 주차장의 차량 출입 관리 등을 위해서는 차량 번호판 인식은 필수적인 과정이다. 자동차 번호판 인식을 위한 기존 대부분의 방법은 번호판의 모양, 질감, 색상 등의 특징을 활용하여 번호판의 위치를 찾고 여러 가지 영상분할 알고리즘을 적용하여 숫자 및 문자영역으로 분할한 뒤 분류(classification)알고리즘을 사용하여 번호판을 인식한다[2-4]. 그러나 일반 도로나 골목 등 개방된 환경에서 촬영된 번호판 영상은 햇빛이나 야간 가로등 조명의 영향과 같은 주변환경 변화가 많이 포함된다. 또한 비 혹은 눈이 내리거나 흐린날과 같은 날씨 변화에 따른 빛의 반사 등으로 차량 번호판에 잡음이 많이 포함되고 색상 변화가 심할 수 있다. 그리고 카메라의 위치 혹은 자동차의 진행 방향에 따라 기하학적인 왜곡현상이 발생할 수도 있다. 따라서 이런 문제를 해결하기 위해 색상보정 혹은 기하학적인 왜곡 보정을 위한 방법을 추가로 사용해야 한다[3]. 또한 칼라영상을 흑백영상으로 변환한 뒤 에지를 찾고 이를 가로 및 세로 방향으로 투영하여 자동차 번호가 포함된 영역을 검출하고 세선화 과정 등을 거쳐 숫자 및 문자를 인식하는 휴리스틱(heuristic) 방법이 있다[3,4].

본 논문에서는 카메라로부터 획득된 원본 자동차 번호판 영상으로부터 번호영역을 추출하기 위해 의미론적(semantic) 영상분할과 딥러닝(deep learning) 알고리즘을 사용한다[5-8]. 의미론적 영상분할은 자율주행 자동차 연구분야에서 자동차 주행중에 주변 환경을 인식하는데 주로 사용하는 알고리즘이다. 본 논문에서는 획득된 번호판 원본 영상에 기존의 화소 투영과 같은 추가적인 전처리과정 없이 딥러닝 및 의미론적 분할을 적용하여 바로 숫자 및 문자영역을 검출하는 방법을 제안한다.

2. 관련연구

자동차 번호판 검출을 위한 대표적인 방법은 컬러 정보를 이용하는 방법, 텍스춰 정보를 이용하는 방법, 에지 및 수학적 모폴로지(morphology) 를 이용하는 방법 등이 있다[3,4,9].

칼라정보를 이용하는 방법은 주로 번호판의 색상이 서로 다르게 되어 있다는 것을 이용한다. 즉 영업용 차량은 노랑색 바탕에 검정색 번호, 일반용 차량은 흰색 바탕에 검정색 번호, 외교용차량은 감청색 바탕에 흰색 번호, 친환경 차량은 하늘색 바탕에 검정색 번호로 되어 있다. Oh[9]등은 번호판에서 숫자의 칼라는 흰색과 검정색으로 제작되며 숫자는 일정한 간격으로 배치된다는 성질을 이용하여 번호판 숫자영역을 추출하였다. 칼라정보를 활용하는 방법은 먼저 번호판 영역을 자동차 영상으로부터 검출한 후 직사각형의 종회비 및 가장자리 밀도와 같은 특징을 기반으로 번호영역을 추출한다. 그러나 이 방법은 번호판과 유사한 색상을 갖는 영역이 있는 경우 번호판 영역을 검출하는데 어려움이 있으며, 잡음 및 조명변화에 따라 결과가 다르게 될 수도 있다. Long[10]등은 조명변화에 강인한 번호판 영역 검출방법을 제안하였는데 HSI 칼라 영역으로 변환한 후 밝기성분에 대하여 소벨연산자로 수직 에지성분을 검출한다. 그리고 길이가 작은 잡음 에지 및 곡선에지 등을 제거하고 남아 있는 에지영상에 대하여 직사각형 창을 스캔하면서 번호판 영역을 검출하는 방법을 제안하였다.

에지 정보를 이용하는 방법은 번호판 영역의 에지 밀도가 상대적으로 높다는 특성을 이용하는 것으로 조명변화 및 잡음에 민감한 단점이 있다[4,11,12]. Duan[14]등은 Hough변환을 이용하여 입력영상에서 강한 에지성분을 추출하고 번호판의 종횡비를 고려하여 번호판 영역을 추출하였다. 그러나 처리시간이 상대적으로 길며 메모리 사용량 및 조명변화 등에 따른 오류가 발생하는 단점이 있다.

3. 자동차 번호영역 검출을 위한 제안방법

3.1 사용된 신경망 모델

본 논문에서는 딥러닝과 의미론적 영상분할을 사용하여 자동차 번호판의 숫자 및 문자영역을 검출하는 방법을 연구하였다. 그림1에 제안방법의 블록도를 표시하였다. 의미론적 영상분할은 화소 단위로 분류된 기준영상(ground-truth)이 필요하다. 본 논문에서는 영상레이블러(image labeler)를 사용하여 번호판 영상을 화소 단위로 레이블링한다. 즉 번호판의 숫자와 문자영역이 포함된 전경부분(foreground)과 이를 제외한 배경부분(background)으로 레이블링하고 이를 딥러닝 신경망 학습에 사용한다.

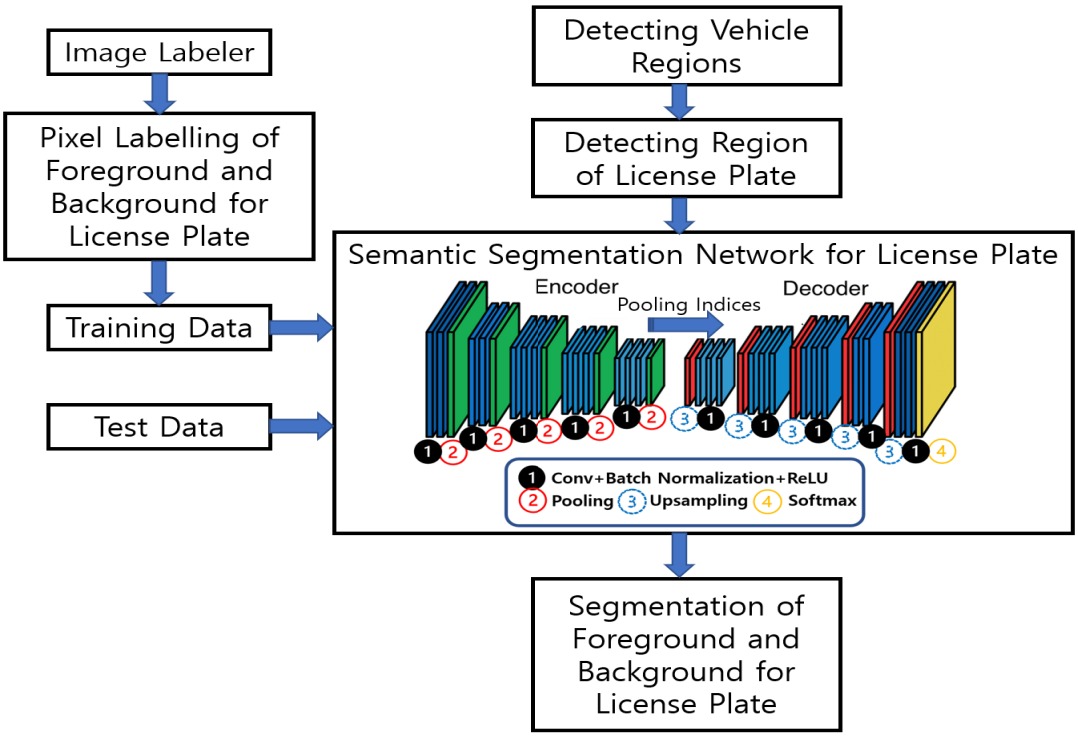

Fig. 1. Block diagram of the proposed method

본 논문에서는 자동차 번호영역을 분할하는 기저모델(base model)로 Vgg계열 및 ResNet계열을 사용한다[7,13,15]. Vgg모델은 옥스퍼드대학 연구팀인 VGG(Visual Geometry Group)에 의해 개발된 신경망으로 이미지 분류를 위한 대표적인 모델로 알려져 있다. Vgg모델은 신경망 은닉층의 수에 따라 표1과 같이 Vgg16 모델과 Vgg19 모델로 나뉜다. Vgg모델은 기존의 신경망과 다르게 특징맵(feature map)을 얻기 위한 필터의 크기를 3x3으로 하였다. 즉 3x3 필터로 영상에 2회 및 3회 합성곱(convolution)을 적용하여 얻은 특징맵은 5x5 및 7x7필터를 각각 2회 및 1회 적용하여 얻은 특징맵과 크기가 같게 된다. 이렇게 하면 학습시켜야 할 모델의 매개변수의 수가 크게 줄어 학습속도를 높일 수 있는 반면에 신경망 은닉층의 수가 증가하게 되어 비선형성이 높아지고 이는 모델의 특징 식별성이 향상되는 결과를 가져온다.

ResNet계열은 마이크로소프트에서 개발한 모델로 2015년 ImageNet Challenge에서 우승을 한 딥러닝 신경망 모델이다. 일반적으로 신경망의 계층이 증가할수록 성능이 우수해지는 것으로 알려져 있다. 그러나 ResNet 개발팀은 신경망의 계층 수가 증가함에 따라 오류 역전파 과정에서 기울기 소실 및 폭발 때문에 성능이 저하됨을 밝혔다. 여기서 기울기 소실은 지도학습에서 오류를 역전파할 때 오류를 편미분한 기울기를 사용함에 따라 역전파 단계가 깊어질수록 미분값이 0에 가깝게 되

Table 1. Structure of Vgg16 and Vgg19

Vgg16	Vgg19	# of feature map
conv3 conv3 maxpool	conv3 conv3 maxpool	64
conv3 conv3 maxpool	conv3 conv3 maxpool	128
conv3 conv3 conv3 maxpool	conv3 conv3 conv3 conv3 maxpool	256
conv3 conv3 conv3 maxpool	conv3 conv3 conv3 conv3 maxpool	512
conv3 conv3 conv3 maxpool	conv3 conv3 conv3 conv3 maxpool	512
fully connected layer-4096/4096/1000		
softmax layer		

어 오류가 전달되지 않는 문제이며, 기울기 폭발은 반대로 기울기가 갑자기 증가하는 현상이다. ResNet 신경망에서는 이를 해결하기 위해 스킵연결(skip connection)를 추가한 잔차학습(residual learning)을 통해 신경망의 계층 수가 증가하여도 기울기 소실 및 폭발 문제가 발생하지 않도록 하였다.

ResNet 신경망은 3x3 합성곱이 반복된다는 점에서는 Vgg 신경망과 구조가 유사하다. 그러나 ResNet은 그림 2와 같이 임의 계층 입력을 출력에 바로 더해주는 스킵연결을 추가하였다. 이렇게 하면 ResNet은 $F(x)+x \simeq x$가 되도록 신경망을 학습하는 것과 같고 이는 결국 $F(x)$가 0이 되도록 학습하는 것이 된다. 여기서 x는 임의 신경망의 입력이고 $F(x)+x$는 출력이다. 따라서 $F(x)$는 잔차(residual)에 해당하고 ResNet은 잔차 $F(x)$가 0이 되도록 학습하게 된다. 그리고 신경망의 은닉층 수에 따라 ResNet18, ResNet50 등으로 구별된다.

3.2 의미론적 영상분할

의미론적 영상분할을 위한 신경망의 내부는 그림1과 같이 부호화(encoder) 및 복호화(decoder) 구조로 되어 있다. 의미론적 분할은 각 화소 단위로 클래스 정보와 위치 정보를 동시에 고려한 영상분할 방법이다.

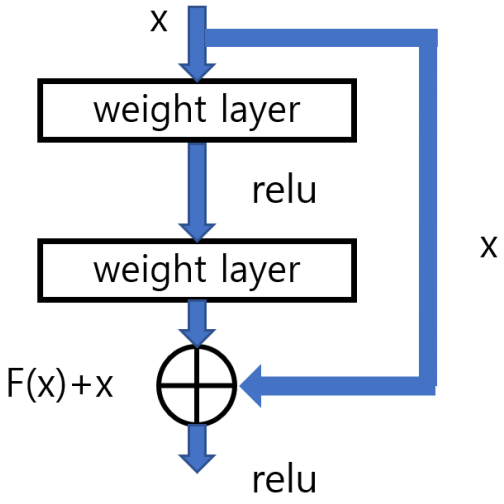

Fig. 2. Connection structure of ResNet

입력영상이 심층신경망에 입력되면 부호화 과정은 그림1과 같이 ❶합성곱, 배치정규화 및 ReLU(rectified linear unit)을 거치고 ②풀링(pooling)을 하여 특징맵의 크기를 줄인다. ReLU 과정은 신경망의 비선형성을 증가시키고 풀링 과정은 위치정보를 유지하면서 특징맵의 차원을 줄이는 효과가 있다. 그리고 부호화 단계의 마지막은 완전연결층(fully connected layer)이 아닌 1x1 합성곱 층을 추가하여 특징맵에서 각 물체의 위치정보를 보존할 수 있게 한다. 즉 의미론적 분할을 위한 신경망은 마지막 특징맵에 클래스 정보뿐만 아니라 위치정보도 포함하고 있는 것이 일반적인 분류신경망과 다른 점이다. 의미론적 분할을 위한 심층신경망은 영상분할을 목적으로 하므로 복호화과정은 부호화 단계의 마지막 특징맵을 그림 1과 같이 풀링인덱스를 넘겨받아 다시 ③upsampling 하여 입력영상과 같은 크기로 만들어야 한다. 그리고 복호화의 마지막층은 미리 정해진 클래스별로 각 화소를 분류해야 하므로 ④소프트맥스(softmax) 층이 필요하다.

4. 실험결과 및 고찰

제안방법의 성능을 평가하기 위해 본 연구에서는 실제 도로 위의 CCTV 카메라에서 촬영된 영상을 사용하였다. 먼저 촬영된 동영상에서 AlexNet 합성곱 신경망을 이용한 전이학습을 통하여 자동차 영역을 분리하고 번호판 영역을 검출하는 선행연구를 진행하였다[16]. 본 논문에서는 후속 연구로 추출된 번호판 영역에 의미론적 분할 방법을 적용하여 숫자 및 문자영역을 검출하는 방법을 제안한다.

실험에 사용된 번호판 영상은 500장으로 하루 중 햇빛과 날씨의 변화에 따른 색상변화가 포함되어 있다. 그림 3은 실험에 사용된 번호판 영상 일부이다. 블러링(blurring) 잡음이 포함되어 있으며 주변 환경 변화로 인해 영상 화질이 선명하지 않음을 알 수 있다.

색상변화에 따른 성능저하를 줄이기 위해 칼라 정규화가 필요하며, 그림 4에 칼라 정규화 및 크기를 조정한 영상을 표시하였다. 그림 5는 제안방법을 적용한 실험결과이다. ⓐ는 원 영상이고 ⓑ는 기준영상, 그리고 ⓒ는 Vgg16를 기저 신경망으로 사용하여 구한 의미론적 분할 영상이다. ⓓ는 원영상과 분할영상을 겹쳐서 표시한 것이고 ⓔ는 기준영상과 분할영상을 겹쳐서 표시한 것으로 ⓓ 및 ⓔ를 통해 원 영상의 번호영역과 추출된 번호영역이 겹치는 공통영역이 높다는 것을 알 수 있다.

Fig. 3. Original license plate

Fig. 4. Color normalized and resized license plate

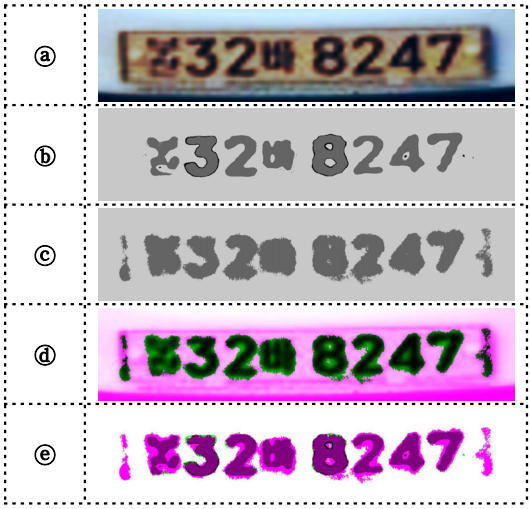

Fig. 5. Original image and experimental results using Vgg16 model. ⓐOriginal image ⓑGround-truth image ⓒSegmented image ⓓOverlap with original and segmented result ⓔOverlap with ground-truth and segmented result

본 논문에서 의미론적 영상분할을 위해 사용한 기저 신경망은 Vgg16, Vgg19, ResNet18 및 ResNet50이다. 각 모델에 따른 실험 결과를 그림 6에 나타내었다. 그림 6의 ⓐ 및 ⓒ는 원영상이고 ⓑ 및 ⓓ는 해당 기준영상의 일부이다. 각 모델별로 성능을 비교하기 쉽도록 기준영상

Fig. 6. Comparison of simulation results according to neural network models. ⓐ,ⓒ Original image ⓑ,ⓓ Partial enlargement of ground-truth image

일부를 확대한 것이다. 녹색부분은 기준영상에서는 어떤 클래스(전경 혹은 배경)에도 포함되지 않았던 화소를 분할영상에서는 전경부분으로 분류된 화소이다. 실험결과 Vgg계열 신경망을 사용할 때는 번호 및 테두리 부분에 잡음이 검출되었으나 ResNet계열은 상대적으로 번호영역 검출성능이 우수함을 알 수 있다.

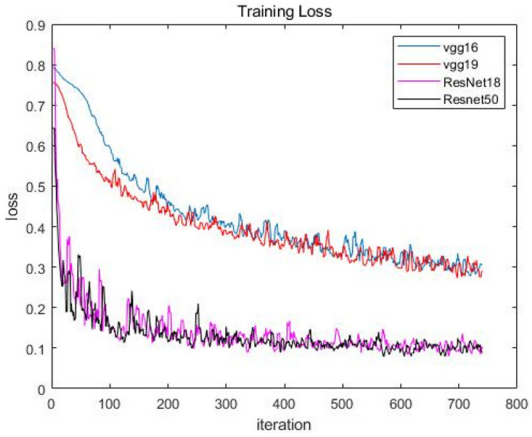

Fig. 7. Convergence characteristics of cross entropy loss function using sgdm optimization

그림 7은 최적화 방법으로 sgdm(stochastic gradient descent with momentum)을 사용할 때 교차엔트로피 손실(loss)함수의 수렴 특성 일부를 나타낸 것으로 ResNet50 신경망이 가장 빠르게 수렴함을 알 수 있다.

Table 2. Accuracy, mean IoU and mean BF value for each neural network model

metric / model	Accuracy(%)	mean IoU	mean BF
Vgg16	90.83	0.7742	0.8798
Vgg19	91.29	0.7822	0.8956
ResNet18	94.65	0.8665	0.9694
ResNet50	95.77	0.8790	0.9722

표 2는 각 신경망 모델에 따른 정확도, 평균 IoU (intersection of union), 평균 BF(boundary F1)값을 표시한 것으로 의미론적 분할의 결과를 비교하는 대표적인 척도(metric)이다. 정확도(accuracy)는 기준영상 대비 분할영상이 얼마나 정확하게 분할되었는가를 나타내는 지표이다. IoU는 기준영상과 분할영상 사이에 겹치는 부분(공통영역)의 화소 수를 두 영역을 합한 전체 화소 수로 나눈 값인데, 평균 IoU는 각 클래스별 IoU를 평균한 것이다. BF는 기준영상과 분할영상의 경계 일치도를 표시하는 지표이다.

본 논문에서 사용한 주요 하이퍼파라미터(hyperparameters)값으로 드롭아웃 비율은 30%, 모멘텀은 0.9, 학습률은 0.001로 하였으며 최적화 방법은 sgdm을 사용하였다. 표 2와 그림 6, 7에서 알 수 있는 바와 같이 ResNet50의 경우 95.77% 정확도를 보여 성능이 가장 우수하며 분할영상에 숫자 및 문자가 대부분 포함된 것을 알 수 있다.

5. 결론

본 연구에서는 기존방법에서 사용하던 에지검출, 수직 및 수평 방향으로의 투영과정과 같은 전처리과정 없이 의미론적 분할을 이용하여 자동차 번호판 영상에서 숫자 및 문자영역을 원 영상으로부터 바로 추출하는 방법을 연구하였다. 조명 및 날씨 변화에 따른 번호판 영상의 색상변화를 줄이기 위해 정규화하였으며 전경 및 배경영역의 면적 비율 차이에 따른 편향을 줄이는 방법을 연구하였다. 의미론적 분할을 위한 심층신경망 학습을 위해 영상레이블러로 기준영상을 생성하였으며, 심층신경망으로 Vgg16, Vgg19, ResNet18과 ResNet50을 사용하여 성능을 비교하였다. 실험결과 번호판에서 숫자 및 문자영역 검출은 ResNet50을 사용할 때 성능이 가장 우수하였다. 향후 연구과제는 제안방법을 전체 교통관제 시스템에 연결하는 연구가 필요하다.

REFERENCES

[1] S. An, B. Lee & D. Shin. (2011).A Survey of Intelligent Transportation Systems. *Third International Conference on Computational Intelligence, Communication Systems and Networks*, Bali, 332-337. DOI: 10.1109/CICSyN.2011.76

[2] M.A. Lalimia, S. Ghofrania & McLernonb. (2013). A Vehicle License Plate Detection Method Using Region and Edge Based Methods. [Online]. DOI: 10.1016/j.compeleceng.2012.09.015.

[3] M. Ko. (2004). *Effective License Plate Character Recognition based on Geometric Invariant Features.* Doctoral dissertation, Kyungpook National University, Daegu.

[4] H. Bai & C. Liu. (2004). A Hybrid License Plate Extraction Method Based on Edge Statistics and Morphology. *Proceedings of the 17th International Conference on Pattern Recognition,* 831-834. DOI: 10.1109/ICPR.2004.1334387

[5] J. Schmidhuber. (2014). *Deep Learning in Neural Networks: An Overview.* [Online]. arxiv. org/abs/1404.7828. DOI:10.1016/j.neunet.2014.09. 003

[6] A. Khan1, A. Sohail1, U. Zahoora1 & A. S. Qureshi1.(2020). *A Survey of the Recent Architectures of Deep Convolutional Neural Networks.* [Online]. arxiv.org/abs/1901. 06032. DOI: 10.1007/s10462-020-09825-6

[7] Vijay Badrinarayanan, Alex Kendall & Roberto CipollaSegNet. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39,* 2481-2495. DOI: 10.1109/TPAMI.2016.2644615

[8] J. Lee. (2018). A Method of Eye and Lip Region Detection using Faster R-CNN in Face Image. *Journal of the Korea Convergence Society, 9(8)*, 1-8. DOI: 10.15207/JKCS.2018.9.8.001

[9] B. Oh & D. Cho. (2011). Number Region Extraction of License Plates Using Colors and Arrangement of Number. *Journal of Korea Multimedia Society*, *14(9),* 1117-1124.

[10] X. Long & J. Zhou. (2013). An Efficient Method for Dark License Plate Detection. *Int. Journal of Digital Content Technology and its Applications*, *7(55),* 477-485. DOI: 10.4156/JDCTA.VOL7.ISSUE1.55

[11] R. Chen & Y. Luo. (2012). An Improved License Plate Location Method Based on Edge Detection. *Proceeding of International Conference on Applied Physics and Industrial Engineering*, 1350-1356.

[12] K. Lim, H. Byun & Y. Cho. (2016). Vehicle License Plate Detection in Road Images. *Journal of Korean Institute of Information Science and Engineers*, *43(2),* 186-195.

[13] K. Simonyan & A. Zisserman. (2015). *Very Deep Convolutional Networks for Large-scale Image Recognition.* [Online]. arXiv:1409.1556. DOI:10.2146/ajhp170251

[14] T. D. Duan, D. A. Duc & T. L. H. Du. (2004). Combining Hough Transform and Contour Algorithm for Detecting Vehiclesí License-Plates. *Proceeding of International Symposium on Intelligent Multimedia, Video and Speech P rocessing*, 747-750.

[15] H. Li & C. Shen. (2016). *Reading Car License Plates Using Deep Convolutional Neural Networks and LSTMs.* [Online]. arXiv:1601.05610

[16] Y. Jeong, I. Ansari, J. Shim & J. Lee. (2017). A Car Plate Area Detection System Using Deep Convolution Neural Network. *Journal of Korea Multimedia Society. 20(8),* 1166-1174. DOI: 10.9717/KMMS.2017.20.8.1166

이 정 환(Jeong-Hwan Lee) [정회원]

· 1983년 2월 : 경북대학교 전자공학과 (공학사)
· 1993년 8월 : KAIST 전기및전자공학과(공학 석·박사)
· 1990년 3월 ~ 현재 : 안동대학교 전자공학과 교수
· 관심분야 : 인공지능, 영상처리, 생체인식, 딥러닝 등
· E-Mail : jhlee@anu.ac.kr