

SSD-Mobilenet과 ResNet을 이용한 모바일 기기용 자동차 번호판 인식시스템

(Vehicle License Plate Recognition System using SSD-Mobilenet and ResNet for Mobile Device)

김윤기*, Fatemeh Dehghan*, 조성원 **

(Woonki Kim, Fatemeh Dehghan, Seongwon Cho)

요약

본 논문은 고성능의 서버 없이 안드로이드 스마트폰 단독으로 동작할 수 있도록 경량화 딥러닝 모델을 사용하여 구현한 자동차 번호판 인식 시스템을 제안한다. 자동차 번호판 인식시스템은 [번호판검출]-[문자영역 분할]-[문자인식]으로 3단계의 과정으로 구성되며, 번호판검출은 SSD-Mobilenet, 문자영역 분할은 ResNet에 localization을 추가하여 사용하였고 문자인식은 ResNet을 이용하여 구현하였다. 테스트한 기기는 삼성 갤럭시 S7, LG Q9이며 정확도는 약 85.3%, 실행속도는 약 1.1초가 소요된다.

■ 중심어 : 자동차 번호판 인식 시스템 ; 딥러닝 ; SSD-Mobilenet ; ResNet

Abstract

This paper proposes a vehicle license plate recognition system using light weight deep learning models without high-end server. The proposed license plate recognition system consists of 3 steps: [license plate detection]-[character area segmentation]-[character recognition]. SSD-Mobilenet was used for license plate detection, ResNet with localization was used for character area segmentation, ResNet was used for character recognition. Experiments using Samsung Galaxy S7 and LG Q9, accuracy showed 85.3% accuracy and around 1.1 second running time.

■ keywords : Vehicle License Plate Recognition ; Deep Learning ; SSD-Mobilenet ; ResNet

I. 서론

기존의 자동차 번호판 인식시스템은 속도위반과 신호 위반 감시, 불법 주정차 감시와 같이 기관의 감시 시스템으로 사용되며, 대개 특정 임베디드 기기용 펌웨어로 구현되거나 카메라로부터 얻은 이미지를 고성능 딥러닝 서버로 보내 인식결과를 받아오는 방식으로 구현된다[1-3].

최근에 안드로이드는 가장 널리 사용되는 운영체제가 되면서, 자동차 번호판 인식시스템을 안드로이드로 구현하는 경우 공개계를 이용한 블랙박스, 중고차량 정보조회, 자율 주행과 같은 서비스로 확장할 수 있으며, 기존의 감시 시스템에서도 공개계 안드로이드 폰에 앱만 설치하면 되기 때문에 유지 보수하기 쉬우며, 휴대성이 좋아 실시간 정보조회 및 적발 등 활용할 수 있다.

임베디드 펌웨어로 LPR(License Plate Recognition)을 구현하는 경우 주로 영상처리 기법을 이용하여 번호판과 문자영역을

찾게 되는데 이는 딥러닝에 비해 속도가 빠르며 메모리가 적게 필요한 장점이 있지만 한 이미지에서 하나의 번호판만 검출할 수 있다는 한계점과 영업용, 외교용 번호판이나 새로 추가된 전기차, 2019년형 번호판같이 다양한 종류의 번호판의 모든 경우를 고려하여 알고리즘을 수정 및 추가하기에 어려움이 크다.

딥러닝 서버와의 통신을 이용해 LPR을 구현하는 경우 학습할 데이터의 양이 충분하다면 여러 차량의 다양한 번호판에 대해서 검출할 수 있으나 기기 측에서는 통신에 사용되는 데이터 비용이, 서버 측에서는 요청량이 과다하게 되는 경우 부하가 확대되는 문제점이 있다.

본 논문에서는, 안드로이드 기반 스마트 모바일 기기 자체에서 동작이 가능하도록 경량화된 딥러닝 모델을 이용하여 구현된 자동차 번호판 인식 시스템을 제안한다.

* 준회원, 홍익대학교 대학원 전자전기공학과

** 정회원, 홍익대학교 전자전기공학부

본 연구는 정부재원(중소벤처기업부) 지원을 받아 수행되었습니다. 연구비 지원에 감사드립니다.

접수일자 : 2020년 02월 07일

수정일자 : 2020년 04월 13일

게재확정일 : 2020년 05월 06일

교신저자 : 조성원, e-mail : swcho@hongik.ac.kr

II. 기존 연구

1. 한국 차량 번호판 종류

한국에서 사용되고 있는 차량 번호판은 2종류의 크기 규격을 가지고 있다. 2006년형 이전에 주로 사용되던 규격은 가로 335mm, 세로 170mm로 종횡비가 약 1:2이며, 2006년 이후에는 가로 520mm, 세로 110mm로 종횡비가 약 1:5인 크기로 변경되었다. 크기 이외에도 목적과 용도에 따라 일반, 구형, 사업용, 친환경, 외교용 등 여러 종류의 차량 번호판이 사용되고 있다.

표 1. 한국 차량 번호판 종류

구분	유럽식 비율(1:5)	미국식 비율(1:2)
일반	28마 7073	1076897
구형	1996년형	서울 46 마 3227
	2004년형	33기 3509
사업용	설 35바 2819	경기 88 사 8768
친환경	20마 9597	
외교용	외교 085-010	외교 001-281

한국 차량 번호판에서 사용 가능한 문자는 다음 표 2와 같이 지정되어 있으며 지역이나 외교같이 의미 단위를 가지는 문자를 1개로 가정할 경우 2019년 현재 최대 9개의 문자열로 구성된다.

표 2. 한국 차량 번호판에 사용 가능한 문자

구분	사용 가능 글자
숫자	0 1 2 3 4 5 6 7 8 9
일반	가 나 다 라 마 거 너 더 러 머 버 서 어 지 고 노 도 로 모 보 소 오 조 구 누 두 루 무 부 수 우 주
지역	서울 부산 인천 광주 대구 대전 울산 경기 강원 충북 충남 전북 전남 경북 경남 제주
택시 / 버스	바 사 아 자
택배	배
렌터카	하 허 호
군용	국 합 육 해 공
외교용	외교 영사 준외 준영 국기

2. 영상처리를 이용한 번호판인식

차량 번호판 인식을 위해서는 다음과 같은 3단계 작업이 필요하다. 먼저 이미지 내에서 자동차 번호판의 위치를 검출한 다음, 각 글자에 해당하는 문자영역을 분할 한 후, 각 문자가 무엇이었는지 인식하는 과정을 거친다.



그림 1. 번호판 인식시스템 과정

영상처리를 이용하여 자동차 번호판을 인식하는 경우, 전체 이미지에서 Canny 에지 알고리즘으로 윤곽선을 검출한 후 수직 방향 에지 연결 및 수평으로 에지 전파하는 LSD(Line Segment Detection)를 통해 번호판 후보를 찾아낸 후 에지 밀도와 종횡비를 통해 최종적인 번호판을 검출한다[5].

검출된 번호판 영역에서 이진화 작업을 수행한 후 CLNF(CCLUF with NFPP) 알고리즘을 통해서 각 문자영역에 대한 정보를 얻을 수 있으며, 번호판과 마찬가지로 각각의 문자영역 또한 종횡비를 이용하여 최종적인 분할된 문자영역을 얻을 수 있다[6].

단순 영상처리를 통해서 번호판검출 및 문자영역을 분할하는 경우 딥러닝보다 속도가 빠르다는 장점이 있으나 조명 환경 및 노이즈에 대해 민감하며, 번호판 종류에 따라 검증에 사용될 알고리즘을 추가로 작성해야 하는 문제점이 있다.

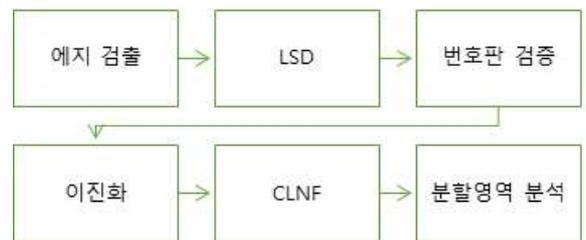


그림 2. 단순 영상처리 기반 LPR 과정

3. Mask R-CNN을 이용한 번호판인식

Mask R-CNN은 자체적으로 2단계의 구성으로, 관심 영역을 제한하는 단계와 그 관심 영역의 물체가 무엇이고, 어디에 있는지를 검출하는 단계로 구성되어 있다[7]. 이 모델을 이용하여

기존의 3단계 구성이 아닌 모델 하나만으로 모든 과정이 가능하나 모델에 사용되는 파라미터가 약 5천만 개로 연산량 또한 매우 많으므로 서버용 컴퓨터로는 연산할 수 있으나 현재의 임베디드 또는 모바일 AP 성능으로는 사용하기에 적합하지 않다.

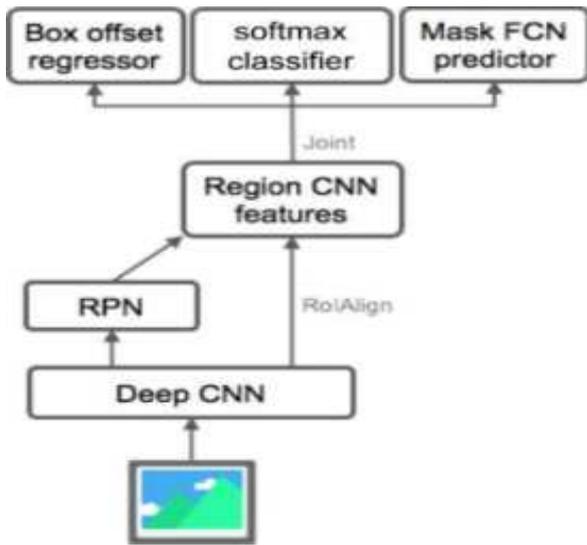


그림 3. Mask R-CNN의 구조도

III. 제안하는 차량 번호판 인식 시스템

제안하는 차량 번호판 인식 시스템은 Mask R-CNN과 같이 무거운 모델을 사용하는 것이 아닌 가벼운 모델들을 각각의 단계별로 최적화하는 것을 목표로 한다.

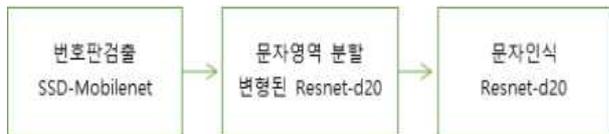


그림 4. 제안하는 차량 번호판 인식 과정

1. 번호판검출 단계

차량 번호판검출 단계는 이미지에서 정해지지 않은 개수의 여러 번호판을 검출해야 하며 크기가 다양할 수 있으므로 가벼운 Detector인 SSD-Mobilenet을 사용하였다.

SSD는 convolution과 pooling레이어에서 가지는 기하학적인 위치정보를 이용하여 특징을 추출하여 앵커를 통해 임의의 위치 및 임의의 크기의 물체를 검출할 수 있도록 하는 네트워크 구성 방식이며, Mobilenet은 depthwise convolution과 pointwise convolution을 이용하여 기존의 convolution보다

커널 크기 만큼 연산량을 줄인 네트워크이다[8-9].

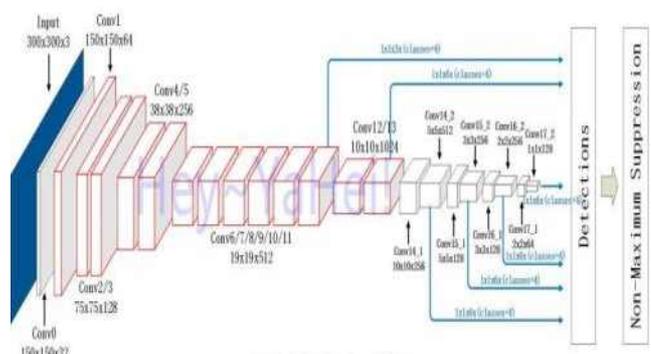


그림 5. SSD-Mobilenet의 모델 구조

SSD-Mobilenet에서 사전 학습된 모델에서 자동차에 대한 특징을 충분히 뽑을 수 있으므로 전이 학습을 이용하였고, 앵커박스 중횡비를 자동차 번호판 비율에 맞게 1:2, 1:5로 최적화 하였다.

표 3. 차량 번호판검출 단계의 입출력

입력 이미지 형태 [300, 300, 3]	
모델 파라미터 개수	6,375,232개
출력 [2268, 72]	앵커: 2268개 앵커 정보: 8 예측 위치정보: 4 클래스: 60 (실제 사용 클래스: 1)

2. 문자영역 분할 단계

문자영역을 분할하는 단계는 앞서 차량 번호판을 검출하는 과정과는 달리 문자의 최대개수가 정해져 있는 특징이 있어, ResNet[10]을 변형하여 최대 9개의 문자를 attention 하도록 구성하여, 각 문자의 위치와 숫자, 한글, 특수문자, 배경을 분류할 수 있도록 구성하였다.

ResNet은 기존의 딥러닝 네트워크에서 레이어가 20계층 이상 쌓아 올릴 경우 목적함수를 통해 계산된 경사 값이 역전파가 될수록 점차 작아져 첫 번째 레이어에 도달할 경우 경사가 사라지는 문제점이 있으나, 이러한 문제를 일정한 간격마다 출력 레이어에 입력 레이어를 더하는 형태로 경사 값에 대한 정보를 잃지 않고 전달할 수 있는 경로가 되어 학습이 잘 안 되는 문제를 해결하였다.

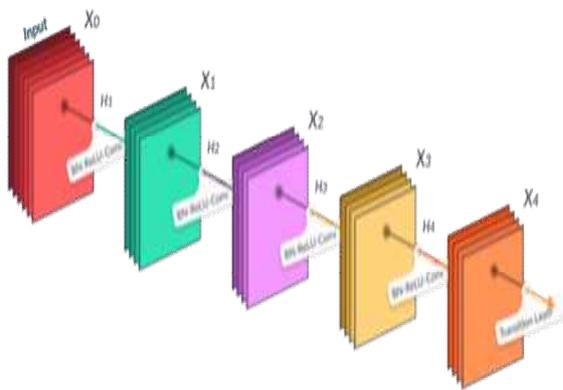


그림 6. ResNet의 모델 구조

ResNet은 기본적으로 분류기이며 위치에 대한 정보를 가지고 있지 않기 때문에 마지막 레이어에 각 문자당 4개의 노드를 추가, mean square error(MSE)를 적용하여 문자를 학습시켰고 이를 변형 ResNet으로 명명하기로 한다.

$$E = \frac{1}{n} \sum_i^n (y_{true} - y_{pred})^2 \quad (1)$$

표 4. 문자영역 분할 단계 입출력 형태

입력 이미지 형태 [w:160, h:32, c:3]	
모델 파라미터 개수	283,172개
출력 [9, 8]	<div style="display: flex; justify-content: space-around;"> 배경 숫자 한글 특수 </div> <div style="display: flex; justify-content: space-around; margin-top: 10px;"> xmin ymin xmax ymax </div> <p>[4 분류, 2개의 점 위치], 최대 9개 박스</p>

3. 문자인식 단계

문자인식 단계에서는 ResNet을 통해 분류를 수행하였으며 이전 단계에서 분류된 결과인 숫자인지 한글인지 특수문자인지에 따라 조건에 따라 argmax를 적용하였고 문자가 부족한 데이터나 기술어진 상황에 대해서는 data augmentation을 적용하여 보완하였다.

표 5. 문자인식 단계 입출력 형태

입력 이미지 형태 [w:32, h:32, c:3]	
모델 파라미터 개수	234,999개
출력 [76]	<div style="display: flex; justify-content: space-around; margin-bottom: 10px;"> 0 1 ... 9 </div> <p style="text-align: right;">: 10개</p> <div style="display: flex; justify-content: space-around; margin-bottom: 10px;"> 가 거 ... 호 </div> <p style="text-align: right;">: 40개</p> <div style="display: flex; justify-content: space-around;"> 서울 인천 ... 외교 </div> <p style="text-align: right;">: 26개</p>

표 6. 각 단계별 사용한 파라미터 개수

모델	파라미터 개수	메모리 크기
번호판검출 (SSD-Mobilenet)	6,375,232	36.6MB
문자영역 분할 (ResNet with attentions)	283,172	2.2MB
문자인식 (ResNet)	234,999	1.8MB
합계	6,893,403	모델: 40.6MB 전체: 196.4MB

휴대기기에서는 하드웨어의 발전에 따라 RAM 용량이 커지는 경향이 있지만 오래된 기기에서도 호환성을 고려한다면 메모리를 적게 차지할수록 유리하므로, 정확도를 유지하되 최대한 파라미터의 수를 줄이는 것이 중요하다.

사용된 딥러닝의 전체 파라미터의 개수는 6,893,403개로 대부분의 파라미터는 SSD-Mobilenet에서 사용하고 있다. 문자영역 분할과 문자인식 단계는 레이어의 수와 필터의 개수를 조정하여 번호판 검출에 비해 훨씬 적은 파라미터 개수로 동작할 수 있도록 최적화 하였다.

IV. 실험 결과

실험에 사용한 휴대기기는 삼성 갤럭시 S7 엣지, LG Q9 으로 테스트하였으며 다른 앱에 의한 영향을 줄이기 위해 공장 초기화를 진행한 후 비행기모드로 전환하여 와이파이, 블루투스, 모바일 데이터 등을 끄고 실험하였다.

실험에 사용한 이미지는 지하주차장 및 남산 3호 터널 톨게이트에서 촬영된 이미지를 사용하여 실험하였다.

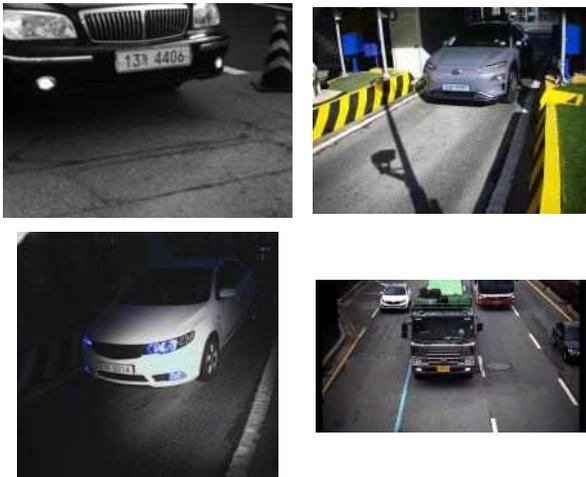


그림 7. 실험에 사용한 이미지

정확도에 대한 기준은 번호판검출에서는 IOU > 0.7을 만족하면서 내부의 모든 문자는 포함할 경우를 기준으로 하였으며, 문자영역 분할에서는 모든 문자에 대해서 IOU > 0.7을 기준으로 하였고, 문자인식에서는 모든 문자가 일치하는 경우이다. IOU(Intersection of union)는 예측 박스와 정답 박스 사이에 합집합에 대한 교집합의 비로 예측 박스와 정답 박스 사이에 얼마나 일치하는지에 대한 지표가 된다.

표 7. 번호판 유형별 정확도 비교

	표본 수	오류 수	정확도
일반(1:5.5)	3826	572	84.8%
일반(1:2)	452	46	89.8%
영업용	30	6	80.0%
전기차	2	0	100.0%
구형	30	10	66.7%
외교	5	1	80.0%
합계	4389	645	85.3%

구형(녹색) 번호판의 경우 1996년형 개정판, 2004년 개정판으로 2가지 타입이 존재하며 1996년형의 경우 학습시킨 표본의 수가 많지 않아 오류 수가 높게 검출되었다. 그 외의 타입에 대해서는 80% 이상의 정확도를 가진다.

오류가 생기는 경우는 초점이 명확하지 않거나, 이미지가 훼손된 경우, 이미지보다 차량이 매우 작게 촬영된 경우, 번호판의 각도가 많이 기울어질수록 오류의 경우가 많이 검출되었다.

표 8. 번호판인식 단계별 정확도 비교

단계	SSD-MobileNet	ResNet (detector)
번호판검출	99.5%	92.2%
문자영역 분할	91.7%	93.0%
문자인식	91.1%	91.7%

표 9. 번호판인식 단계별 속도 비교

단계	SSD-MobileNet	ResNet (detector)
번호판검출	0.863s	0.225s
문자영역 분할	0.832s	0.143s
문자인식	0.850s	0.114s

번호판검출의 경우 속도는 ResNet detector가 빠르나 정확도를 고려하면 SSD-MobileNet을 사용하였고, 문자영역 분할과 문자인식은 속도와 정확도 모두 ResNet (detector)가 우수하여 ResNet (detector)를 사용하였다.

표 10. 제안하는 번호판인식 모델과 정확도와 속도

단계	모델	정확도	실행속도
번호판검출	SSD-MobileNet	99.5%	0.863s
문자영역 분할	ResNet detector	93.0%	0.143s
문자인식	ResNet	91.7%	0.114s
전체		85.3%	1.121s

제안하는 시스템은 전체적으로 85.3%의 정확도를 보이며 실행속도는 평균적으로 1.1초 정도 소요된다. 메모리는 상주 9MB, 최대 30MB만큼 요구한다.

다음 표 11은 다른 차량 번호판 인식 시스템들과 성능 비교 결과이다. 제안 시스템은 모바일 기기로 적용 가능한 모델 중에서 가장 높은 정확도와 빠른 실행속도를 보여준다.

표 11. 다른 시스템과의 성능 비교

구분		정확도	실행속도
모바일 기기	제안 시스템	85.3%	0.03s
적용 가능 차량 번호판 시스템	영상처리 LSD + DBN [5][11]	52.7%	0.02s
	YOLO [3][12]	74.2%	0.11s

V. 결론

본 논문에서는 휴대기기에서 경량화된 딥러닝으로 동작하는 차량 번호판인식 시스템을 제안하였다. 제안된 시스템은 3 단계로 구성되며, 번호판검출은 SSD-MobileNet, 문자영역 분할은 ResNet detector, 문자인식은 ResNet을 적용하여 구현하였다. 실험 결과 다른 방식에 비하여 상대적으로 높은 정확도와 빠른 실행속도로 동작 가능하다는 것을 확인하였다.

REFERENCES

- [1] 송영준, 안재형, 김남, “자동차 번호판 인식 기술 동향,” *정보통신연구진흥원 학술정보(IITA)*, 2006
- [2] M. A. Rafique, W. Pedrycz, and M. Jeon, “Vehicle license plate detection using region-based convolutional neural networks,” *Soft Computing*, June 2017.
- [3] R.-C. Hendry, Rung-Ching Chen, “Automatic License Plate Recognition via sliding-window darknet-YOLO deep learning,” *Image and Vision Computing*, vol. 87, pp. 47–56, July 2019.
- [4] Mohanad Hazim Nsaif Al-Mayyahi, Nawaf Hazim Barnouti, and Mohammed Abomaali, “Vehicle Detection and License Plate Recognition System,” *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 3170–3174, 2018.
- [5] R. G. von Gioi, et al, “LSD:A Fast Line Segment Detector with a False Detection Control,” *IEEE PAMI*, vol. 32, no. 4, pp. 722–732, Apr. 2010.
- [6] 웅웬탄빈, 정선태, “주변 전경 픽셀 전과 알고리즘 기반 실시간 이동 객체 검출,” *전자공학회논문지*, 제41권, 2004년 1월
- [7] K. He, G. Gkioxari, P. Dollár and R. Girshick, “Mask R-CNN,” *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, pp. 2980–2988, 2017.
- [8] Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott E. Reed, Cheng-Yang Fu and Alexander C. Berg, “SSD: Single Shot MultiBox Detector,” *ECCV* (2016).
- [9] Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto and Hartwig Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” *ArXiv abs/1704.04861* (2017).
- [10] K. He, X. Zhang, S. Ren and J. Sun, “Deep Residual Learning for Image Recognition,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, pp. 770–778, 2016.
- [11] 연우, “딥러닝(DBN) 기반 임베디드 차량 번호판 인식 시스템에 관한 연구,” *홍익대학교 석사학위 논문*, 2017
- [12] Redmon, Joseph et al., “You Only Look Once: Unified, Real-Time Object Detection,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.

저 자 소 개



김윤기(준회원)

2017년 홍익대학교 전자전기공학부
학사 졸업.
현재 홍익대학교 대학원 전자
전기공학과 석사과정.

<주관심분야 : 인공지능>



Fatemeh Dehghan(준회원)

2015년 이란 Yazd 대학교 학사 졸업.
현재 홍익대학교 대학원 전자전기
공학과 석사과정.

<주관심분야 : 인공지능>



조성원(정회원)

1982년 서울대학교 학사 졸업.
1987년 Purdue University 석사
졸업.
1992년 Purdue University 박사
졸업.
현재 홍익대 전자전기공학과 교수.

<주관심분야 : 신호처리, 인공지능>