Journal of KIIT. Vol. 17, No. 11, pp. 115-120, Nov. 30, 2019. pISSN 1598-8619, eISSN 2093-7571 115
http://dx.doi.org/10.14801/jkiit.2019.17.11.115

ROI 추출을 위한 기계 학습 기반 영상 전처리 기법

조명원*, 김남기**[1], 이병대**[2]

An Image Preprocessing Scheme Based on Machine Learning for ROI Extraction

Myoungwon Jo*, Namgi Kim**[1], and Byoung-Dai Lee**[2]

본 논문은 2019학년도 경기대학교 대학원 연구원장학생 장학금 지원에 의하여 수행되었음. 또한 2017년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원(No. 2017-0-00794, VR 가상현실 콘텐츠 무선전송을 위한 모듈 개발)과 교육부의 재원으로 한국연구재단의 지원(NRF-2017R1D1A1B04027874)을 받아 수행된 연구임

요 약

왼손 X-ray 영상을 이용한 골연령 측정에서 주로 사용하는 TW3(Tanner - Whitehouse 3)기법은 왼손에서 13개의 ROI(Region Of Interests)를 추출하고 각 ROI의 등급을 판정한다. 그런데 판독 시간이 상대적으로 길고 판단하는 사람에 따라 정확도가 달라지는 문제점이 있다. 이를 해결하기 위해 딥러닝을 이용한 자동화 시스템이 개발되고 있는데 이 시스템의 판독 성능은 ROI 추출 정확도에 의해 크게 좌우된다. 그러나 왼손 X-ray 영상은 영상의 밝기(brightness)와 대비(contrast)에 따라 정확한 ROI 추출 성공률이 달라진다. 따라서 본 논문에서는 기계 학습을 통해 X-ray 영상의 밝기와 대비를 자동으로 조절하여 ROI를 정확히 추출할 수 있는 기법을 제안한다. 제안하는 기법을 사용하면 골연령 측정 자동화 시스템에 부담이 없으면서도 ROI 추출 성공률이 29.5% 정도 더 높아짐을 실험을 통해 알 수 있었다.

Abstract

TW3 (Tanner - Whitehouse 3) method using the left-hand X-ray image is performed by extracting 13 regions of interests (ROIs) from the left hand and determining the grade of each ROI. However, the assessment time is relatively long and the accuracy varies depending on the judge. To solve this problem, an automated system using deep learning is being developed and the performance of the system is greatly constrained by the accuracy of ROI extraction. However, in the left hand X-ray image, the success rate of accurate ROI extraction depends on the brightness and contrast of the image. Therefore, in this paper, we propose a scheme to accurately extract ROI by automatically adjusting the brightness and contrast of X-ray images by machine learning. Experiments show that the proposed scheme increases ROI extraction success rate up to 29.5% without burdening the automated bone age assessment system.

Keywords
image preprocessing, deep learning, brightness, contrast, bone age

* 경기대학교 컴퓨터과학과 석사과정
- ORCID: http://orcid.org/0000-0001-7593-3073
** 경기대학교 컴퓨터과학과 교수
- ORCID[1]: http://orcid.org/0000-0002-0077-6576
- ORCID[2]: http://orcid.org/0000-0002-4028-6168

· Received: Oct. 21, 2019, Revised: Nov. 04, 2019, Accepted: Nov. 07, 2019
· Corresponding Author: Byung-Dai Lee
Division of Computer Science and Engineering, Kyonggi University, 154-42, Gwanggyosan-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, Korea
Tel.: +82-31-249-9676, Email: blee@kyonggi.ac.kr

Ⅰ. 서 론

성조숙증을 비롯한 다양한 질병을 진단하기 위하여 여러 가지 진단법이 있다. 그 중 골연령 검사는 환자의 왼손 X-Ray 영상을 통해 뼈의 모양을 보고 전문의가 등급을 판정하여 골연령을 결정하는 진단법이다. 골연령 검사에서는 GP(Greulich-Plye)기법[1]과 TW3(Tanner - Whitehouse 3)기법[2]이 가장 널리 사용되어지고 있다. GP 기법은 단순하고 시간이 적게 걸린다는 장점이 있지만 골연령 평가를 하는데 정확도가 떨어진다는 단점이 있다[3]. TW3 기법은 GP에 비하여 긴 시간을 필요로 하지만 좀 더 세분화된 단위로 판독이 이루어진다. 그러나 TW3 기법은 판독에 있어서 전문가에 따라 수분에서 수십 분까지 시간이 걸리는 단점이 있다. 이러한 단점을 없애고 전문가의 숙련도에 의한 평가 결과의 편차를 줄이기 위한 딥러닝을 활용한 TW3 기반 골연령 판독 자동화 시스템에 대한 연구가 진행되었다[4]-[6].

딥러닝은 다양한 분야에서 활용되고 있다[7][8]. 이러한 딥러닝을 활용한 TW3 기반 골연령 자동 판독 시스템은 왼손 X-ray 영상을 이용한다. 왼손 X-ray 영상에서 ROI(Region Of Interests) 13개를 추출하고 추출된 ROI 마다 등급을 부여한다. 그리고 부여한 등급의 합을 이용하여 최종 골연령을 판정하게 된다. 이러한 과정에서 딥러닝 기술을 이용하기 위해서는 13개의 ROI를 자동으로 추출하는 것이 매우 중요하다. 그런데 이러한 ROI 추출은 영상에서 뼈와 배경의 밝기(Brightness) 및 대비(Contrast)에 따라 그 성능이 매우 크게 차이난다. 따라서 본 논문에서는 영상 전처리 과정에서 13개의 ROI를 가장 잘 추출할 수 있는 왼손 X-Ray 영상의 밝기 조절 파라미터를 기계학습을 통해 찾는다. 그리고 이 결과를 딥러닝을 활용한 TW3 기반 골연령 측정 시스템에 도입하여 그 성능이 얼마나 좋아지는지 평가한다.

Ⅱ. TW3 기반 골연령 자동 판독 시스템

TW3는 환자의 왼손에서 13개의 검사 부위의 생김새를 정의된 기준과 비교하여 A부터 I까지의 골등급을 부여하고, 각 부위의 골등급에 해당되는 점수를 합산하여 최종적인 점수를 얻게 된다. 최종점수는 TW3에서 성별에 따라 정의한 0.1세 단위의 기준표로 분류되며, 이를 통해 대상자의 골연령을 예측할 수 있다. 그림 1은 TW3 기법에서 사용하는 13개의 ROI를 표시한 것이다.

딥러닝을 이용한 TW3 기반 골연령 자동 판독 시스템은 이러한 TW3 골연령 판독 방법을 딥러닝으로 학습하여 자동화한 시스템이다[4]. 골연령 자동 판독 시스템에서 가장 중요한 것은 13개의 ROI를 정확히 추출하고 각 ROI의 등급을 정확히 판단하는 것이다. 특히 13개의 ROI를 정확히 추출하는 것은 올바른 판독을 위해서는 첫 번째로 중요한 단계라 할 수 있다. 그런데 그림 1에서 보다시피 왼손 뼈에서 ROI 영역과 검지와 약지 부분에 비슷한 모양을 가진 다른 영역이 다수 존재한다.

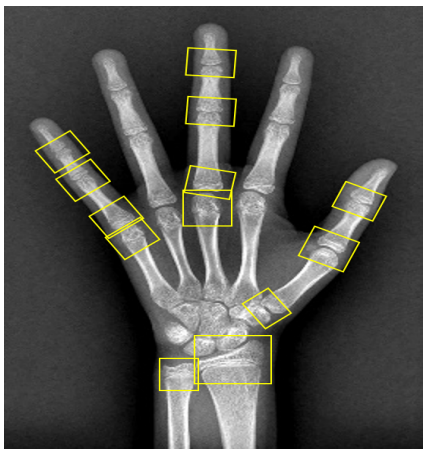


그림 1. TW3 기법에서 13개의 ROIs
Fig. 1. 13 ROIs in TW3 scheme

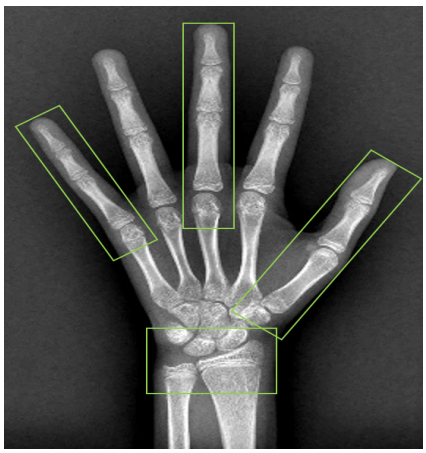


그림 2. 4개의 BigROIs
Fig. 2. 4 BigROIs

그러므로 본 논문에서 소개하는 골연령 자동 판독 시스템은 영상에서 13개의 ROI를 직접 추출하지 않고 더 큰 영역인 엄지, 중지, 소지, 손목 BigROI를 먼저 추출하고 각 영역 안에서 13개의 ROI를 추가 추출하는 방법을 사용하였다. 이렇게 되면 비슷한 생김새로 인한 잘못된 ROI 추출을 획기적으로 줄일 수 있다. 그림 2는 왼손에서 네 개의 BigROI를 나타낸 것이다.

그림 3은 딥러닝을 이용한 TW3 기반 골연령 자동 판독 시스템의 전체 구조이다. 해당 시스템에서 제안한 첫 번째 과정은 영상 전처리 기법으로 왼손 X-Ray 영상에서 BigROI 영역을 추출한다. 그리고 추출한 각각의 BigROI 영역에서 딥러닝 기반 Faster-R-CNN 객체 검출 기법[9]을 이용하여 13개의 세부 ROI를 추출한다.

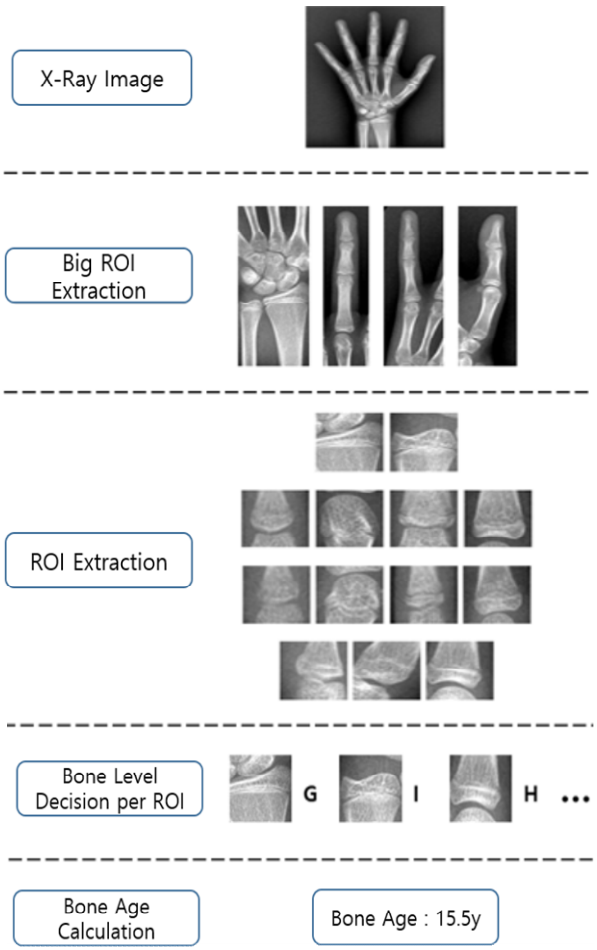


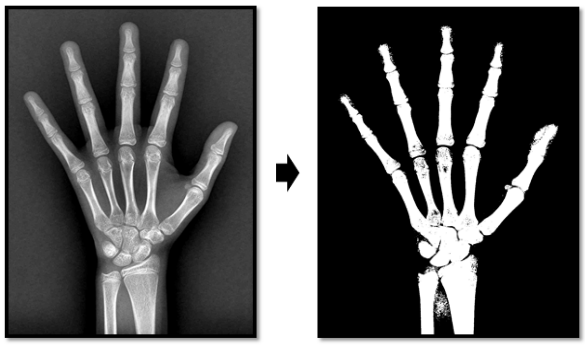
그림 3. 골연력 자동 판독 시스템 구조
Fig. 3. Architecture of bone age automatic assessment system

세 번째 과정은 추출된 13개의 ROI의 각각의 골등급을 분류하는 과정이다. 이 과정에서는 딥러닝 기법 중에 영상 분류를 하는데 있어서 많이 사용되어지는 CNN을 사용한다. 학습을 하는데 있어서 필요한 데이터는 원본 X-Ray 영상에서 수작업으로 추출한 13개의 ROI와 전문의가 판독한 각 ROI의 골등급이 필요하다. 의료 영상의 경우 데이터 수집을 하는데 있어 어려움이 있기에 데이터 증가 과정을 거쳐 1개의 영상을 총 13개로 증가시킨다. 기존 CNN에서 사용되어지는 신경망구조인 VGGnet을 기반으로 골 영상의 특징을 더욱 잘 학습 할 수 있게 수정한 VGG-BAnet을 사용한다. 그리고 데이터가 한정적이기에 10 Fold Cross Validation을 사용하여 모델의 신뢰도를 높인다. 각 ROI는 각각의 10개의 ROI로부터 등급을 판독하고 10개의 결과를 앙상블(Ensemble)하여 골등급을 최종적으로 결정한다. 마지막으로는 위 과정에서 얻은 13개의 골등급을 점수로 계산하여 RUS Score를 확인하여 TW3 기법에서 제시된 기준으로 골연령을 판독한다.

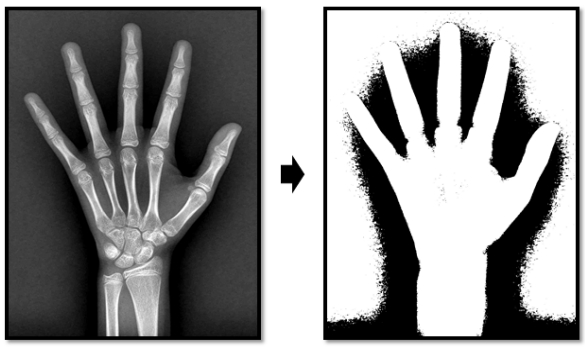
III. 자동 영상 전처리 기법

앞서 소개한 TW3 기반 골연령 자동 판독 시스템에서 어떤 X-ray 영상이던 관계없이 성공적으로 BigROI를 잘 추출하기 위해서는 X-ray 영상의 밝기와 대비를 적절히 조절하는 영상 전처리 과정이 반드시 필요하다. 골연령 자동 판독 시스템에서는 Big ROI 추출을 위해 뼈 부분과 배경 부분을 분리하기 위한 이진화 과정을 거친다.

그림 4는 이러한 이진화 과정의 예를 보여 주고 있다. 그림에서 알 수 있듯이 이진화 과정에서 영상의 밝기와 대비 정도는 이진화 결과에 큰 영향을 미친다. 그림 4(a)는 영상이 적절한 밝기와 대비를 가지고 있어서 뼈와 배경이 올바르게 이진화된 예를 보여 주고 있다. 반면 그림 4(b)는 영상의 배경 부분이 너무 밝아 이진화 과정에서 뼈와 배경이 제대로 분리되지 않는 예를 보여 주고 있다. 따라서 영상의 상태에 따라 적절히 밝기와 대비를 조절해 줄 수 있는 방법이 반드시 필요하다.



(a) 잘 이진화된 예
(a) A good example of binarization



(b) 잘못 이진화된 예
(b) A bad example of binarization
그림 4. 왼손 X-Ray 영상 이진화 예
Fig. 4. An example of left-hand X-ray image binarization

BigROI 추출 과정에서 손과 배경을 보다 정확하기 분리하기 위해 본 논문에서는 원본 X-Ray 영상의 밝기와 대비를 딥러닝을 통해 자동으로 조절할 수 있는 방안에 대해 제안한다. 영상을 조절할 때는 밝기를 조절하는 값 α와 대비를 조절하는 값 β를 필요로 한다. 영상마다의 필요한 α와 β 값이 다르기 때문에 본 논문에서는 딥러닝을 활용하여 α와 β 값을 얻어낸다. 제안하는 방식은 영상 전처리 과정에서 영상의 각 픽셀들의 RGB 값을 읽어와 밝기를 조절하는 임의의 변수 α와 대비를 조절하는 임의의 변수 β를 가지고 각 픽셀 P를 식 (1)을 통해 계산하여 P'으로 변경한다. 그리고 변경된 픽셀로 이미지를 전처리한 후 골연령 자동 판독 시스템의 입력으로 사용한다.

$$P_R^{'} = \alpha \times P_R + \beta \qquad (1)$$
$$P_G^{'} = \alpha \times P_G + \beta$$
$$P_B^{'} = \alpha \times P_B + \beta$$

Ⅳ. 실 험

운영체제를 비롯한 소프트웨어 및 하드웨어 환경은 표 1과 같다. 표에서 보듯이 운영체제는 리눅스 계열의 Ubuntu를 사용했고 프로그래밍 언어는 Python, 딥러닝 라이브러리는 Tensorflow를 사용하였다. 또한 하드웨어 GPU는 GeForece 1080Ti를 사용하였다.

표 1. 실험환경
Table 1. Experimental environments

Type	Name
OS	Ubuntu 1.604[10]
Language	Python 2.7[11]
Deep learning library	Tensorflow 1.5[12]
CUDA	CUDA 9.0[13]
Cudnn	Cudnn 7.1[14]
GPU	GeForce 1080Ti[15]

영상 전처리 자동화를 위해 딥러닝의 경우 간단한 Fully Connected Network를 사용하였으며, 모델을 만들고 테스트하기 위해 사용된 학습 데이터 집합은 1,000장이고 검증 데이터 집합의 경우 배경의 밝기가 밝아서 기존의 시스템에서 실패 할 수도 있다고 예상되는 데이터를 포함하여 남녀 100장씩 총 200장으로 정하였다.

영상 전처리 시에 필요로 하는 값은 밝기를 조절하기 위한 α와 대비를 조절하기 위한 β 이렇게 두 개다. 따라서 α를 위한 모델 Mα를 학습하고 β를 위한 모델 Mβ을 따로 만들어 기존 영상 전처리 과정에 추가하였다. 이를 보다 자세히 설명하면 다음과 같다. 골연령 자동 판독 시스템에서 먼저 외손 X-Ray 영상을 입력받는다. 입력받은 왼손 X-Ray 영상을 Mα를 사용하여 밝기 조절을 위한 α을 얻고, Mβ를 사용하여 대비 조절을 위해 필요로 하는 β를 얻는다. 그 후에 해당 값들을 사용하여 영상 전처리를 통해 밝기와 대비를 조절하여 영상을 가공한다. 그 후에는 기존의 딥러닝 기반 골연령 자동 판독 시스템과 동일하게 진행된다. 이렇게 영상 전처리를 통해 한번 가공된 영상은 BigROI를 찾는 작업을 진행할 때 손 영역을 찾기 위한 이진화 작업에서 손과 배경의 분리를 정확하게 할 수 있게 된다.

표 2는 영상 전처리 과정을 추가하기 전과 추가한 후의 BigROI 추출 성공율을 보여 주고 있다. 표에서 알 수 있듯이 본 논문에서 제안하는 영상 전처리 과정을 추가할 경우 기존 시스템에 비하여 남성 X-Ray 영상의 경우 21% 증가하였으며 여성 X-Ray 영상의 경우 약 38% 증가하였다. 전체적으로 봤을 때는 약 29.5% 증가 증가한 것을 알 수 있다. 따라서 기존 시스템에 실패하던 경우들이 본 영상 전처리 과정을 추가할 경우 성공적으로 개선됨을 알 수 있었다.

표 2. 제안하는 영상 전처리 기법 적용 전/후 성공률
Table 2. Success rate before and after adapting the proposed image preprocessing

		Before	After
Exact BigROI extration rate	Male	0.79	1.00
	Female	0.62	1.00
	Total	0.705	1.00

표 3은 영상 전처리 적용 이전과 이후에 따른 골연령 판독 결과의 정확도를 보여 주고 있다. 표에서 영상 전처리 과정을 추가하기 전후의 RMSE와 MAE의 값의 변화는 거의 없음을 알 수 있다. 따라서 영상 전처리 과정으로 인한 시스템 성능 저하는 없다고 판단할 수 있다.

표 3. 제안하는 영상 전처리 적용 전후에 따른 RMSE, MAE
Table 3. TRMSE and MAE according to before and after adapting the proposed image preprocessing

	Before		After	
	RMSE	MAE	RMSE	MAE
doctor 1	0.677	0.509	0.668	0.508
doctor 2	0.718	0.548	0.720	0.545

표 4는 각 작업 진행 과정별 소요 시간을 나타낸 것이다. 표에서 알 수 있듯이 영상 전처리에 소요되는 평균 시간은 전체 시스템 진행 평균 시간 25.099초에 비해 0.239초로 매우 짧음을 알 수 있다. 그리고 기존의 동일한 과정의 경우 영상 전처리를 추가하기 전후에 시간차이가 거의 없었고 영상 전처리를 적용한 전체 작업 소요시간을 비교했을 때도 아주 미미한 차이만이 있었다. 따라서 영상 전처리로 인한 시스템의 추가 지연 시간은 그 크기가 미미하다고 판단할 수 있다.

표 4. 각 시스템 당 작업별 평균 소요 시간 (초)
Table 4. Average time per task (second)

Processing	Before	After
Image preprocessing	-	0.239
Big ROI	0.235	0.224
ROI	8.722	8.957
Level estimation	15.54	15.55
Bone age calculation	0.132	0.129
Total	24.629	25.099

V. 결 론

본 논문에서는 TW3 기법 기반 골연령 자동 판독 시스템에서 보다 정확하게 ROI를 자동으로 찾기 위해 기계학습을 기반으로 하는 자동화된 영상 전처리 기법을 제안하였다. 제안한 기법을 사용하면 왼손 X-ray 영상에서 전처리를 통해 영상의 밝기와 대비를 조절하여 TW3 기반 골연령 측정을 위한 ROI 영역을 보다 정확히 찾을 수 있다. 또 제안한 기법 적용시 TW3 기반 골연령 측정을 위해 필요한 ROI 영역을 29.5% 정도 더 잘 찾을 수 있음을 실험을 통해 알 수 있었다. 그리고 이러한 전처리 과정이 ROI별 골연령 등급 판정 정확성에는 거의 영향을 미치지 않고, 영상 전처리에 드는 시간도 전체 판독 시간에 비해 매우 작아 자동 판독 시스템에 부담이 크게 없음을 알 수 있었다.

향후에는 단순한 영상 전처리 과정에서 기계학습 방법을 Fully-Connect Network가 아니라 GAN (Generative Adversarial Networks)을 사용하여 보다 나은 성능을 낼 수 있는지 평가해 볼 예정이다.

References

[1] W. W. Greulich and S. I. Pyle, "Radiograph atlas of skeletal development of the hand and wrist", California Stanford University Press, 2nd Ed., 1959.

[2] J. M. Tanner, "Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 method)",

Elsevier, 3rd Ed., 2001.
[3] J. F. Griffith, J. C. Y. Cheng, and E. Wong, "Are western skeletal age standards applicable to the Hong Kong Chinese population? A comparison of the Greulich and Pyle method and the Tanner and Whitehouse method", Hong Kong Medical Journal, Vol. 13, No. 1, pp. 28-32, Feb. 2007.
[4] S. J. Son et al., "TW3-Based Fully Automated Bone Age Assessment System Using Deep Neural Networks", IEEE Access, Vol. 7, pp. 33346-33358, Mar. 2019.
[5] G. Litjens, "A Survey on Deep Learning in Medical Image Analysis", Elsevier Medical Image Analysis, Vol. 42, pp. 60-88, Dec. 2017.
[6] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv preprint arXiv:1409.1556, Sep. 2014.
[7] K. M. Lee and C. H. Lin, "Video Stabilization Algorithm of Shaking image using Deep Learning", Journal of IIBC, Vol. 19, No. 1, pp. 145-152, Feb. 2019.
[8] J. W. Kim and P. K. Lee "Image Recognition based on Adaptive Deep Learnin", Journal of IIBC, Vol. 18, No. 1, pp. 113-117, Feb. 2018.
[9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 39, No. 6, pp. 1137-1149, Jun. 2017.
[10] https://ubuntu.com/ [accessed: Oct. 1, 2019]
[11] https://www.python.org/download/releases/2.7/ [accessed: 1 Oct. 2019]
[12] https://www.tensorflow.org/ [accessed: Oct. 1, 2019]
[13] https://developer.nvidia.com/cuda-90-download-archive [accessed: Oct. 1, 2019]
[14] https://developer.nvidia.com/cudnn [accessed: Oct. 1, 2019]
[15] https://www.geforce.com/hardware/desktop-gpus/geforce-gtx-1080-ti/specifications [accessed: Oct. 1, 2019]

저자소개

조 명 원 (Myoungwon Jo)

2018년 2월 : 경기대학교 컴퓨터과학과(공학사)
2018년 3월 ~ 현재 : 경기대학교 컴퓨터과학과 석사과정
관심분야 : 네트워크, IoT, 딥러닝

김 남 기 (Namgi Kim)

1997년 2월 : 서강대학교 컴퓨터학과(공학사)
2000년 2월 : KAIST 전산학과 (공학석사)
2005년 2월 : KAIST 전산학과 (공학박사)
2005년 3월 ~ 2007년 2월 : 삼성전자 통신연구소 책임연구원
2007년 3월 ~ 현재 : 경기대학교 컴퓨터과학과 교수
관심분야 : 통신 시스템, 네트워크

이 병 대 (Byoung-Dai Lee)

1996년 2월 : 연세대학교 전산학과
1998년 2월 : 연세대학교 전산학과 (공학석사)
2003년 2월 : Univ. of Minnesota CSE (공학박사)
2003년 3월 ~ 2010년 2월 : 삼성전자 책임연구원
2010년 3월 ~ 현재 : 경기대학교 컴퓨터과학과 교수
관심분야 : 기계학습, 딥러닝, 의료 영상 분석