

딥러닝을 이용한 WTCI 설태량 평가를 위한 유효성 검증

이우범*

상지대학교 정보통신공학과

An Effectiveness Verification for Evaluating the Amount of WTCI Tongue Coating Using Deep Learning

Woo-Beom Lee*

Department of Information Communication Engineering, Sangji University

요 약 한방 설진에서 WTCI(Winkel Tongue Coating Index) 설태 평가는 환자의 설태량 측정을 위한 중요한 객관적인 지표 중의 하나이다. 그러나 이전의 WTCI 설태 평가는 혀영상으로부터 설태 부분을 추출하여 전체 혀 영역에서 추출된 설태 영역의 비율을 정량적으로 측정하는 방법이 대부분으로 혀영상의 촬영 조건이나 설태 인식 성능에 의해서 비객관적 측정의 문제점이 있었다. 따라서 본 논문에서는 빅데이터를 기반으로 하는 인공지능의 딥러닝 방법을 적용하여 설태량을 분류하여 평가하는 딥러닝 기반의 WTCI 평가 방법을 제안하고 검증한다. 설태 평가 방법에 있어서 딥러닝의 유효성 검증을 위해서는 CNN을 학습 모델로 사용하여 소태, 박태, 후태의 3가지 유형의 설태량을 분류한다. 설태 샘플 영상을 학습 및 검증 데이터로 구축하여 CNN 기반의 딥러닝 모델로 학습한 결과 96.7%의 설태량 분류 정확성을 보였다.

• 주제어 : 설진, WTCI, 설태량, 딥러닝, 콘볼루션 뉴럴 네트워크

Abstract A WTCI is an important criteria for evaluating an amount of patient's tongue coating in tongue diagnosis. However, Previous WTCI tongue coating evaluation methods is a most of quantitatively measuring ration of the extracted tongue coating region and tongue body region, which has a non-objective measurement problem occurring by exposure conditions of tongue image or the recognition performance of tongue coating. Therefore, a WTCI based on deep learning is proposed for classifying an amount of tonger coating in this paper. This is applying the AI deep learning method using big data. to WTCI for evaluating an amount of tonger coating. In order to verify the effectiveness performance of the deep learning in tongue coating evaluating method, we classify the 3 types class(no coating, some coating, intense coating) of an amount of tongue coating by using CNN model. As a results by testing a building the tongue coating sample images for learning and verification of CNN model, proposed method is showed 96.7% with respect to the accuracy of classifying an amount of tongue coating.

• Key Words : Tongue Diagnosis, WTCI, the amount of Tongue Coating, Deep Learning, Convolutional Neural Network

Received 11 December 2019, Revised 22 December 2019, Accepted 26 December 2019

* Corresponding Author Woobeom Lee, Department of Information&Communication Engineering, Sangji University, 83, Sangjidae-gil, Woonju-city, Gangwon-do, Korea. E-mail: beomlee@sangji.ac.kr

I. 서론

현재 생명, 국방, 기계, 농업, 의학 등의 모든 분야에서 인공지능(AI; Artificial Intelligence) 기술을 활용하기 위한 연구가 활발히 진행되고 있다. 특히 의료 분야에서는 진단의 정확성을 위한 임상적 진료 데이터 분석 및 의료 영상 인식뿐만 아니라, 신약 개발을 위한 임상 실험에서도 시간과 비용의 감소를 목적으로 인공지능 기술을 적용하고 있다.

이러한 인공지능 기술 가운데에서 고전적 신경회로망 모델에서 학습의 효과를 높이기 위하여 많은 층과 층 내의 많은 뉴런으로 구성된 네트워크를 사용하는 딥러닝(Deep learning) 기술은 사진, 동영상, 음성 정보를 학습하여 분류하고 예측하는 기술로서 구글, 페이스북, MS, 트위터 등의 기업에서도 그 성능을 인정하여 다양한 연구가 수행되고 있다.

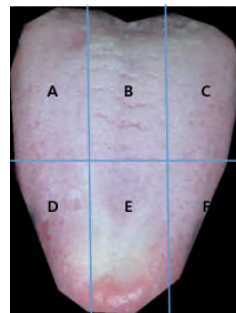
1998년 LeCun 등에 의해서 제안된 콘볼루션 신경회로망(CNN; Convolutional Neural Network)[1]은 주어진 영상을 해석하기 위하여 영상 전체를 해석하는 것이 아니라, 영상처리의 필터링 방법을 기반으로 영상의 특정 부분에 대해서 선택적으로 반응하는 신경생리학적 해석을 참조한 선택적 주의(Selective attention) 모델로서 시각적 영상 분석에 적합하다. 즉 여러 개의 콘볼루션 층으로 구성되어 있으며, 각 층은 적용하는 필터의 크기와 가중치 형태에 따라서 고유의 형태 특징을 추출하여 다음 층으로 전송한다. 이때 CNN 구조에서의 학습은 필터의 결정에 있어서 직관적이나 반복적 실험에 의한 통계적 데이터를 이용하는 것이 아니라 각 단계에서 특징 추출을 위해서 사용되는 필터를 영상 분류의 정확도를 향상시키도록 학습하는 과정을 의미한다. 이러한 CNN 구조를 기반으로 영상 인식의 성능 향상을 위한 다양한 연구가 진행되어 왔으며, AlexNet[2], ZFNet[3], VGG Net[4], GoogLeNet[5], Microsoft ResNet[6]이 대표적이다.

따라서 본 논문에서는 한방 설진에서 환자의 질병 진단이나 건강 상태를 판별하기 위하여 중요한 지표로 사용되고 있는 설태량 측정을 위하여 딥러닝 기술을 적용하고자 한다. 특히 설태량 측정을 위하여 사용되는 WTCI(Winkel Tongue Coating Index) 방법은 혀영상으로부터 추출되는 설태의 영역에 의해서 측정되는 영상 처리 기반의 방법이 대부분으로 혀영상 측정 시의 조명, 촬영 각도, 혀 영역 추출 성능 등의 외부적인 요

인에 의해서 객관성이 떨어지는 문제점이 발생한다. 본 논문에서 이러한 비객관적 요소의 제거를 위하여 CNN 구조의 딥러닝 방법을 적용하여 설태를 소태, 박태, 후태의 3가지 유형으로 학습시키고 학습된 네트워크를 이용한 설태량 분류 인식률을 평가함으로써 WTCI 설태량 평가 방법에 딥러닝 적용의 유효성을 검증한다.

II. WTCI 설태 평가 방법[9]

설진을 위한 설태량 평가를 위해서 본 논문에서 사용하는 설태 평가 방법인 WTCI는 Winkel 등이 제안한 방법으로 그림 1과 같이 혀의 설침부에서 설근부까지 나타나도록 촬영한 혀 전체 영상으로부터 총으로 2등분, 횡으로 3등분하여 총 6구역으로 설태량을 평가한다. 각 구역은 설태 정도에 따라서 설태가 없는 소태의 경우(0점), 설태가 약간 있는 박태의 경우(1점), 설태가 많은 후태의 경우(2점)로 점수를 부여하고 A부터 F까지의 각 구역에 부여된 점수를 합산하여 0~12까지의 측도를 계산하는 방법이다[9].



- 점수/영역
소태 = 0
박태 = 1
후태 = 2
- 점수범위
(0 ~ 12)

$$W.T.C. \text{ Index} = A+B+C+D+E+F$$

Fig. 1. Winkel Tongue Coating Index
no coating(0), some coating(1), intense coating(2)

우리의 이전 연구에서는 설태량의 보다 객관적인 측정을 위해서 WTCI의 값을 3단계 등급으로 부여하는 대신에 각 구역의 전체 면적에서 추출된 설태의 비율을 이용하였다[7, 8]. 그러나 이 방법은 혀영상으로부터 혀몸체(tongue body) 영역과 설태 영역의 추출 성능이 중요한 요소로 작용하기 때문에 한의사의 조작을 요구하는 반수동적 방법으로 사용되었다. 따라서 본 논문

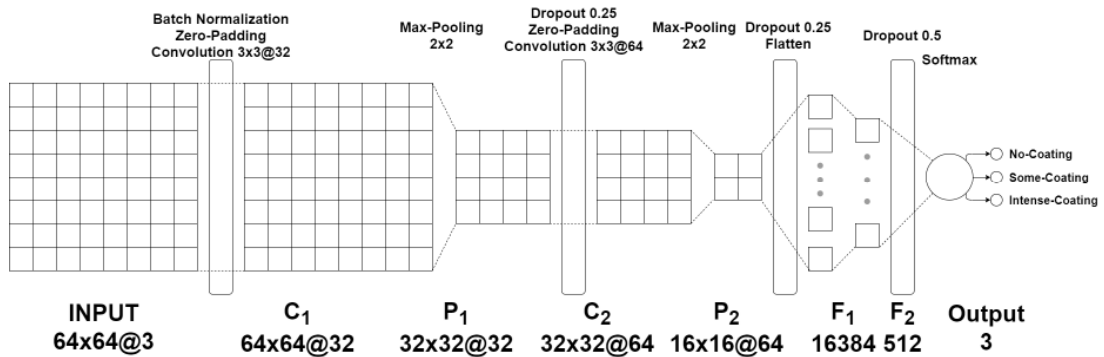


Fig. 2. An CNN Architecture for the WTCI Tongue Coating Evaluation

에서는 한의사의 개입을 최소한으로 하면서 고전적 WTCI 설태 평가에 부합하는 딥러닝 기반의 설태량 유형 인식 방법을 제안한다.

III. 설태량 인식을 위한 딥러닝 구조

설진에서 WTCI 설태량 평가를 위한 CNN 구조는 그림 2와 같다. 신경망의 입력으로는 64x64 화소 크기의 RGB 3채널 혀영상 데이터를 생성하여 사용한다. 그리고 특정 크기를 갖는 필터를 이용하여 샘플 영상으로부터 공간 특징을 추출하기 위하여 C_1 , C_2 두 개의 컨볼루션 층을 구성한다. 두 개의 컨볼루션 층 사이에는 특징 추출 과정에서 배경에 대한 강인함과 공간 내의 특징 성분의 크기에 의한 불변 특징을 획득하기 위하여 P_1 , P_2 두 개의 최대-풀링(Max-pooling) 층을 이용하여 하위-샘플링(Sub-sampling) 연산을 수행한다[1].

초기 입력으로 사용되는 샘플 영상의 경우에 샘플의 촬영 상태에 의한 특징 분포의 반응도 차이를 해소하기 위해 정규화(normalization) 과정을 삽입한다. 정규화 과정은 입력 샘플 영상의 전체 평균과 분산을 이용하여 민감하게 반응하는 특징을 약화하고, 미약한 반응에 대한 특징을 강화한다. 또한 컨볼루션 연산과 사용하는 특징 필터 크기에 의해서 결과 영상의 축소와 손실되는 특징 맵의 가장자리 정보를 보존하기 위해 제로-패딩(zero-padding) 과정을 수행한다[1, 2].

특정 필터에 의한 컨볼루션 연산 후에 2차원 공간 특징 정보가 추출되면 2차원 형태의 다수의 특징 맵을 1차원 형태로 변형하기 위한 플래튼(flatten) 과정을 적용한다. 제안하는 설태량 인식을 위한 CNN 구조에서는 F_1 , F_2 두 개의 완전연결 계층을 사용하여 중간층은

512개 노드로 구성되어 있으며 ReLU 함수를 활성화 함수로 사용한다. 마지막으로 최종 설태량 분류는 소프트 맥스(softmax) 연산을 사용하여 소태, 박태, 후태의 세 가지 설태량 유형으로 인식하여 출력한다.

CNN 학습 과정에서는 과적합 문제를 해결하기 위해 노드 간 연결을 임의적으로 단절하는 드랍아웃(dropout) 과정을 적용한다. 본 논문에서 구성된 CNN의 경우에서 사용하는 드랍아웃 과정은 $\{P_1 \rightarrow C_2; 0.25\}$, $\{P_2 \rightarrow F_1; 0.25\}$, $\{F_2 \rightarrow \text{Output}; 0.5\}$ 의 세 계층 간의 전달 과정에 삽입되어 있다.

IV. 시뮬레이션

설태량 평가에 있어서 딥러닝 적용의 유효성 검증을 위한 CNN 시뮬레이션은 Intel Xeon E5-2620v4, 16G RAM, Geforce GTX1080 8G의 하드웨어 스펙에 Ubuntu 16.04 LTS 운영체제 환경 하에서 오픈 소스 Apache MXNet 1.1.0 프레임워크와 Deep Learning Studio 2.5 플랫폼을 사용하였다[10, 11].

설태 유형 학습을 위한 훈련 데이터 셋은 혀 전체 영상으로부터 64x64 화소 크기의 샘플을 그림 3과 같이 소태, 박태, 후태의 세 가지 유형으로 생성한다. 그림에서 (a)는 소태 샘플로서 전체 샘플 영역에서 1/3 이하 정도의 설태 영역을 포함하며, (b)는 박태 샘플 영상으로 1/3~2/3 정도의 설태 영역, 그리고 (c)는 후태로서 전체 영역의 2/3 이상의 설태 영역을 포함하는 샘플을 의미한다. 설태량 유형 인식을 위하여 CNN 학습 및 검증에 사용된 전체 데이터셋은 총 150개로 설태 유형별로 각각 50개씩 구성하였다. 구축된 데이터 셋으로부터 학습과 테스트를 위하여 전체 150개의 데

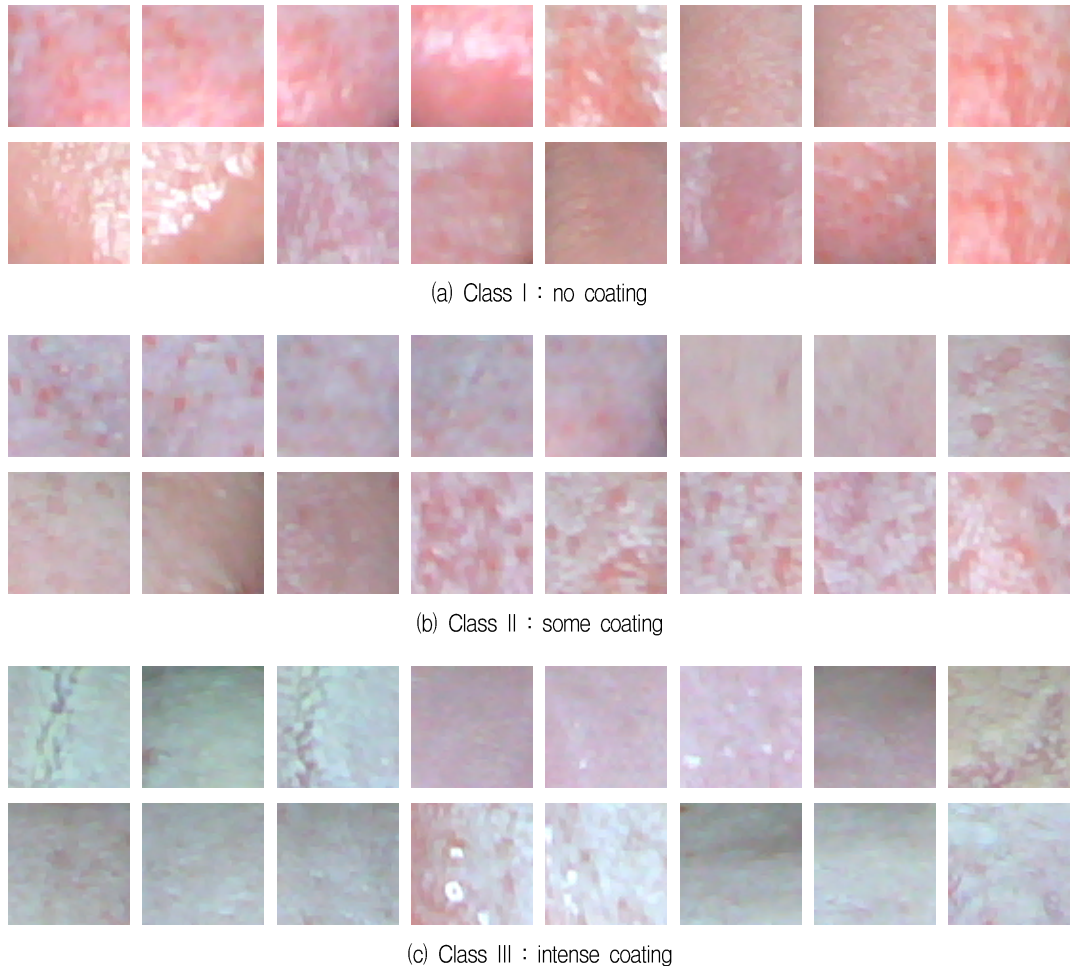


Fig. 3. Example of Sample Dataset for Learning an amount of Tongue Coating

이더넷을 랜덤하게 섞은 후에 임의로 8:2의 비율 정도로 구분하여 120개는 학습 용도로 사용하였고 30장은 유형 인식 성능 테스트를 위한 검증 샘플로 사용하였다.

본 논문에서는 설태량 유형 인식을 위하여 비교적 적은 훈련 데이터셋을 사용하기 때문에 학습과 테스트 검증에 사용하는 훈련 샘플은 각 설태 유형에 따라서 랜덤하게 선택하고 학습률, 모멘텀, 다양한 학습 인수를 설정하여 유효성 검증을 실험하였다. 그림 4의 (a)와 (b)는 총 50 epoch를 갖는 학습 과정에서 발생하는 정확성과 손실 함수의 변화량의 예로서 총 200회의 학습이 진행될수록 정확도는 1로, 손실 함수는 0으로 수렴됨을 나타낸다. 표 1은 학습 후 검증 데이터셋을 이용하여 설태량을 소태, 박태, 후태로 분류한 예시로서

Rating은 테스트 샘플의 부류를 의미하며 Predictions은 CNN에 의해서 추론한 결과 값을 나타낸다. 그리고 Probabilities는 Categorical cross-entropy에 의한 추론 정확도를 확률로 나타내고 있다. 그리고 표에서 추론 결과로 나타난 ‘No’는 소태, ‘some’은 박태, ‘Intense’는 후태를 의미한다. 표 1의 결과에서 정인식의 경우에 90% 이상의 확률로 판단하고 있으며 오인식의 경우에도 90%의 판단율을 나타내고 있다. 이러한 확률은 Cross-entropy에 의해서 계산되는 인식률에 큰 영향을 미치며 그림 4의 정확성과 손실함수 그래프에서도 큰 폭의 변화를 유발함을 알 수 있다.

설태량 분류에 있어서 CNN 적용을 위한 유효성 검증용 학습 데이터셋 30장에 대한 설태량 부류 추론 결과는 표 2와 같다. 표 2에서 CNN은 학습하지 않은 설

태 영상 30장에 대해 96.7%의 추론 정답률을 보였다.

Table 1. Example of Classification of an amount of Tongue Coating











| Image | Rating | Predictions | Probabilities |
|---|---------|-------------|---------------|
|  | No | No | 1.0 |
|  | Intense | Intense | 0.9971 |
|  | No | No | 1.0 |
|  | Some | Some | 0.947 |
|  | No | No | 0.9972 |
|  | Intense | Intense | 0.9997 |
|  | Some | Some | 0.8516 |
|  | Some | Some | 0.9909 |
|  | Some | Some | 0.9675 |
|  | No | Some | 0.9194 |

Table 2. Class Prediction Results of an amount of Tongue Coating

| Rating | predic. | Intense | Some | No |
|----------|-------------------|---------|------|----|
| | Intense | 5 | 0 | 0 |
| Some | 0 | 14 | 0 | |
| No | 0 | 1 | 10 | |
| Accuracy | 96.7%/30(samples) | | | |

이 중 박태와 후태의 경우 100%의 정답률을 보였고, 설태량이 전체 면적의 1/3 정도인 소태 경우 11장 중 1장을 제외하고는 정확한 설태량 부류를 나타냈다. 오 분류 인식된 1장의 경우도 실제 사용된 샘플의 색상이 박태의 색상과 유사성이 크며 육안으로도 박태로 판단 될 가능성이 큰 것으로 평가되었다. 따라서 실험 결과 소태, 박태, 후태 3가지 유형의 설태량 추론을 위한 CNN의 적용은 실험적으로 그 유효성을 보였다.

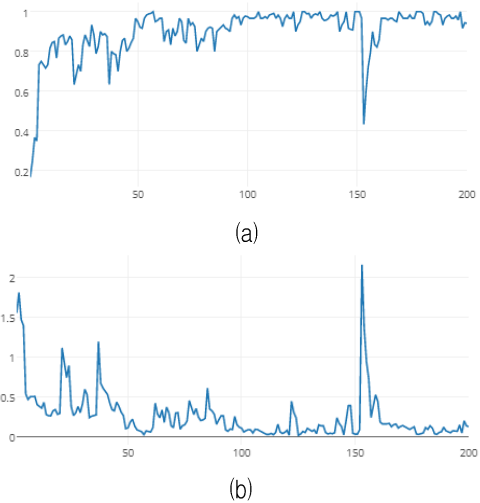


Fig. 4. A validation amount of (a) Accuracy and (b) Loss function in Learning Process.

In case: epoch(50), batch size(30)

loss function(categorical cross-entropy), optimizer(SGD), momentum(0.9), learning rate(0.01)

V. 결론

본 논문에서는 한방 설진에서 설태량 측정을 위한 WTCI 설태 평가에 CNN 기반의 딥러닝 모델을 적용하기 위한 유효성 검증을 수행하였다. WTCI 설태량 평가는 혀영역을 소태, 박태, 후태의 비교적 적은 유형으로 분류하여 적용 가능하기 때문에 빅데이터를 기반으로 하는 영상인식 방법의 적용이 가능하다. CNN 모델을 세 가지 유형의 설태량 평가를 위해서 적용한 결과 비교적 적은 훈련 데이터에도 불구하고 높은 인식률을 나타냈으며, 이것은 화소 비율 기반의 기존의 방법에 비해서 실제 혀영상에서 획득 가능한 샘플 영상의 학습을 기반으로 하기 때문에 설태를 포함하는 혀영상의

가변성에 매우 강하다고 할 수 있다. 그러나 실제 제안한 CNN 모델을 활용하기 위해서는 실제 혀 전체 영상에 대한 구획된 6영역의 통합 인식에 의한 WTCI 평가 모델로의 확장과 샘플 데이터의 확보, 유효성 검증을 위한 평가 모델에 관한 연구가 요구된다.

향후 기존의 WTCI 설태 평가 시스템과의 시스템 수준에서의 딥러닝 인식 모델과의 연동 문제와 딥러닝 모델의 확률적 결과 값의 활용 등의 문제에 관한 연구가 수행된다면 설진에서 설태량 평가를 위한 또 하나의 정량적 평가 방안으로 큰 활용을 기대할 수 있다.

ACKNOWLEDGMENTS

이 논문은 2018년도 상지대학교 교내 연구비 지원에 의한 것임.

REFERENCES

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, 86(11):2278-2324, 1998.

[2] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks.", *advances in neural information processing systems*, pp.1097-1105. 2012.

[3] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional neural networks", *European conference on computer vision*. Springer, Cham, pp.818-833, 2014

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", In *ICLR*, 2015.

[5] Christian Szegedy, et al., "Going Deeper With Convolutions", *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1-9, 2015.

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition", *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.770-778, 2016.

[7] W. B. Lee, "WTCI Tongue Coating Evaluation by analyzing a Ultraviolet Rays Tongue Image Channel",

Journal of The Institute of Signal Processing and Systems, vol. 16(3), pp.96-101, 2015.

[8] W. B. Lee, "Implementation of Computerized Assistant Diagnosis Software for Tongue Diagnosis in the Oriental Medicines", *Journal of The Institute of Electronics and Information Engineers*, vol. 51(6), pp.175-182, 2014.

[9] E. G. Winkel, et al., "Clinical effects of a new mouthrinse containing chlorhexidine, cetylpyridinium chloride and zinc-lactate on oral halitosis", *Journal of Clinical Periodontology*, vol.30(4), pp.300-306, 2003.

[10] *Deep Learning Made Easy with Deep Cognition*, www.kdnuggets.com, 2018.

[11] *Deep Learning Studio Solutions(Exxact)*, www.exxactcorp.com, 2018.

저자 소개

이 우 범 (Woo-Beom Lee)



1995년 : 영남대학교 컴퓨터공학과 학사
 1997년 : 영남대학교 컴퓨터공학과 석사
 2000년 : 영남대학교 컴퓨터공학과 박사
 2000년 3월 ~ 2004년 2월 : 대구과학대학교 컴퓨터공학과 교수
 2004년 3월 ~ 2007년 2월 : 영남대학교 전자정보공학부 컴퓨터전공 특임교수
 2007년 3월 ~ 현재 상지대학교 정보통신공학과 교수
 관심분야 : 뉴로비전컴퓨팅, IT융합의료기(한방의료영상), 딥러닝(신경회로망)