

모바일 환경에서의 심층학습을 통한 심장 순환기 생체신호 분석 진단 도움 기술

이 재 구*, 김 현 재*, 김 성 원*, 송 종 윤*, 윤 성 로^o

Deep Learning-Based Biological Signal Analysis for Assisting Cardiovascular Disease Diagnosis on Mobile Environment

Jaekoo Lee*, Hyunjae Kim*, Sungwon Kim*, Jongyoon Song*, Sungroh Yoon^o

요 약

인체는 다양한 생체기관으로 구성되어 있으며, 각 기관들은 독특한 생체신호들을 발생하며 기능을 수행한다. 수많은 생체신호들 중에서 소리는 전문적으로 훈련된 의료진에 의해 해당 기관의 이상 유무 혹은 질병 진단의 중요한 정보가 된다. 우리는 심장 순환기 청진음을 분석할 수 있는 Deep Learning 기반의 생체신호 분석 진단 도움 기술을 제안하였고, 공개된 심음 데이터를 통해서 방법의 타당성을 확인하였다. 결과적으로 제안한 방법은 대표적인 생체신호인 심음의 Deep Learning 분석 활용에 대한 가능성을 확인하였을 뿐만 아니라 향후, 스마트폰과 같은 모바일 환경에서의 의료 진단 도움 기술 발전에 중요한 연구 토대가 될 것으로 예상된다.

Key Words : Biological Signal, Disease Diagnosis, Deep Learning, Cardiovascular, Mobile

ABSTRACT

Human body consists of various organic systems, and each of these generates peculiar biological signals when it functions. In particular, biological sound involves important diagnostic information of relevant organ, which is manually analyzed by professionally well-trained medical staff. In this paper, we proposed deep learning-based biological sound analysis for assisting cardiovascular disease diagnosis. We confirmed the validity of our method through published real cardiac sound dataset. As a result, we convinced potentiality of applying deep learning method to biological signal-based diagnosis for cardiovascular disease. Furthermore, we expect our work to be an important foundation for development of medical diagnostic system on mobile environment.

I. 서 론 (Introduction)

일반적으로 사람의 뇌, 심장, 폐와 같은 생체기관은 EEG^[1]와 같은 다양한 생체신호들을 발생시킨다^[2]. 특

히, 심음 혹은 호흡음처럼 생체기관에서 발생하는 소리들, 즉 생체소리들은 오랜 시간동안 전문적으로 훈련된 의료진의 청진을 통해 해당 기관의 상태 및 정상 혹은 비정상 여부를 확인하는 중요한 기초 단서가 된

* 본 연구는 미래창조과학부의 재원으로 한국연구재단 바이오의료기술개발사업의 지원 (No.NRF-2014M3A9E2064434, 2015M3A9A7029735) 으로 수행되었습니다.

♦ First Author : Seoul National University Dept. of Electrical and Computer Engineering, jaekolee@snu.ac.kr, 학생회원

° Corresponding Author : Seoul National University Dept. of Electrical and Computer Engineering, sryoon@snu.ac.kr, 종신회원

* Seoul National University Dept. of Electrical and Computer Engineering

논문번호 : KICS2017-03-090, Received March 31, 2017; Revised June 7, 2017; Accepted June 7, 2017

다. 하지만 청진은 의료진의 경험에 따라 질병 판단에 영향을 주는 한계점이 존재한다^{2,3}.

우리는 이런 제한적 상황을 기계학습, 특히 Deep Learning (심층학습)⁴을 통해 해결하려는 방법을 제안하고자 한다. 구체적으로 최근 AI뿐만 아니라 다양한 영역에서 눈에 띄는 결과들을 만들고 있는 Deep Learning 모델을 심음에 적용함으로써 데이터로부터 자동 추출된 특징들 (Data Driven Features)에 기초한 새로운 진단 도움 방법을 제안하고 앞으로의 모바일 환경에서의 의료 진단 도움 기술에 대한 새로운 방향을 모색하고자 한다.

이러한 방법은 실질적인 의료진에게 질병 진단의 편의성을 제공할 뿐만 아니라 모바일 환경에서 일반인에게도 편리하고 간편하게 건강관리를 도와 줄 수 있는 차세대 의료 진단 응용의 핵심기술을 제공할 수 있을 것으로 예상된다. 또한, 다양한 생체신호들을 분석하고 활용할 수 있는 방법을 모색할 수 있는 연구 토대와 방향을 제공하며, 심음을 포함하는 포괄적이고 대규모의 생체신호 수집에 대한 유용성과 이를 Deep Learning 같이 최신의 기계학습 방법들과의 연계 분석에 대한 타당성과 이점도 보여준다.

본 논문은 다음과 같은 순서로 작성되었다. 2장에서는 제안하는 방법에 대한 포괄적인 이해를 돕기 위해 심음을 포함한 일반적인 청진의 일반적인 특성을 설명하고, 이를 분석하는 심층학습의 기본적인 구조와 주요 특징들을 간략하게 제시한다. 서술된 관련 배경 지식들을 토대로 3장에서는 실질적으로 제안하는 방법을 전체에서 세부적인 설명으로 구체화하는 하향식 구조로 제시하고자 한다. 4장에서는 이를 공개된 데이터에 적용한 결과들을 제시함으로써 방법의 타당성을 제공한다. 5장은 제안한 방법에 대한 얻어진 사실들로부터 고려될 수 있는 고찰들과 향후 연구에 대한 방향성을 공유하며, 6장을 통해 우리의 연구 성과를 정리하면서 논문을 마무리한다.

II. 배경 지식 (Background)

2.1 청진 (Auscultation)

청진은 호흡기 및 순환기를 포함한 인체에서 발생하는 다양한 생체소리를 청취하여 관련 기관의 상태와 이상 유무를 판단하는 과정이다^{2,3}. 일반적으로 청취되는 다양한 생체소리는 (그림 1)에 제시된 것처럼 관련 기관의 구조와 구성에 따라 다양한 주파수 대역과 크기를 가지고 있다². 제안한 방법도 주어진 심음의 고유 주파수 대역과 청력 역치 특성을 이용하여 전

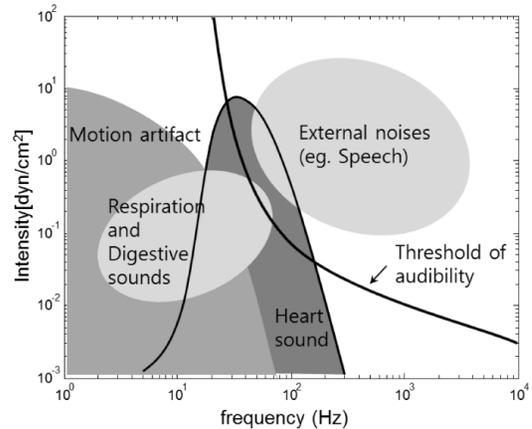


그림 1. 심장 청진음의 주파수 분포와 청력 역치 곡선²
Fig. 1. Frequency distribution of heart sound and hearing threshold curve²

처리하였다.

특히 분석 대상인 심음의 경우, 심장 내 외류에 의해 생겨나는 소리를 측정하여 심장 판막의 협착 및 불완전한 폐쇄, 또는 심장 내의 비정상적인 통로 생성으로 인해 정상적인 혈류가 방해 받아 생기는 병리적 잡음을 통해 심장 질환 여부 판단한다². 심음은 시기 (Timing), 형상적 특징 (Shape), 부위 (Location), 전파/방사 (Radiation), 크기 (Intensity), 고저 (Pitch), 잡음 특징 (Quality) 등의 특성에 따라 분류되며, 이를 토대로 전문적으로 훈련된 의료진은 심장의 정상 혹은 비정상 여부를 판단한다².

2.2 심층학습 (Deep Learning)

인간의 신경망을 모방하여 고차원의 학습을 가능하도록 발전해온 인공신경망은 과거 단순한 퍼셉트론 (Perceptron) 모델로는 제한적인 응용 분야에서만 적용되었지만⁵ 최근 심층학습 모델은 음성 영상인식과 같은 다양한 영역에 적용 및 응용되며, 놀라운 결과들을 만들고 있다⁴. 특히, 잘 알려진 Deep Learning은 기계학습의 한 영역으로 선형과 비선형 변형을 반복적으로 다계층 쌓은 그래프 모델을 이용하여 데이터를 고차원 추상화함으로써 주어진 임무에 맞도록 학습, 추론하는 방법을 의미한다⁴.

제안 방법도 영상 같은 2차원 구조의 입력 데이터 처리에 적합한 특성을 가진 Deep Learning 모델인 합성곱 신경망 (Convolutional Neural Network; CNN)^{4,6}에 기초하고 있으며, (그림 2)의 CNN의 기본 구조를 다층으로 쌓은 모델을 활용함으로써 심음 판단에 도움을 줄 수 있는 유의미한 특징들을 자동적

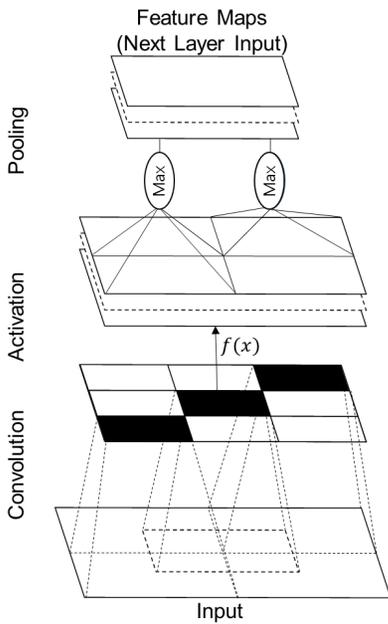


그림 2. Convolutional Neural Network 기본 구조^[4,6]
 Fig. 2. Fundamental structure of Convolutional Neural Network^[4,6]

으로 추출할 수 있었다.

(수식 1) 은 기본적인 CNN의 k 층으로부터 다음 층의 j 번째 합성곱 출력의 연산을 표현하고 있으며, 제안한 방법은 실제 주어진 심음을 2차원 구조로 변환하고 이를 해당 연산을 수행함으로써 CNN에 입력하여 이상 유무를 판단하도록 학습 및 추론을 수행하였다.

$$x_{k+1,j} = h\left(\sum_{i=1}^{\pi_k} F_{k,i,j} * x_{k,i}\right), j = 1, \dots, \pi_k \quad (1)$$

여기서 π_k 는 k층 특징맵 (Feature Map)의 개수이

며, x_k 는 k층의 입력 혹은 이전 층의 특징들이다. 또한 h 는 시그모이드 함수 (Sigmoid Function) 같은 비선형 함수이며, $F_{k,i,j}$ 는 k층의 i번째 특징과 k+1층의 j번째 특징 사이에 실제 학습되는 필터들을 의미한다.

III. 제안 방법 (Proposed Method)

(그림 3)에 제시된 개괄적인 제안 방법을 통하여 우리는 주어진 청취된 심음 데이터에 대해서 최종적으로 심음의 정상 혹은 비정상 여부를 분류 혹은 판단할 수 있는 심층학습 기반의 진단 도움 알고리즘을 제안한다. 구체적으로 제안된 방법은 크게 전처리, 학습, 추정 세 부분으로 나눌 수 있으며, 각 세부 과정들을 순차적으로 설명함으로써 제시된 방법에 대한 이해를 전달하고자 한다.

3.1 전처리 (Preprocessing)

우선, 분석 대상인 심음의 주요 소리 정보가 300Hz 이하의 주파수 대역에서 발생된다는 사실에 기초하여 20-850Hz 대역의 대역 통과 필터 (Band Pass Filter) 를 이용하여 주요 심음 대역의 소리 정보만 추출하여 사용하였다^[2,7]. 또한 대역 내에 존재하는 잡음을 추가 제거하기 위해 Wavelet 변환된 결과의 큰 값을 제거하는 정규화 (Normalization)를 함으로써 잡음 제거 과정을 수행하고^[8], Stein's Unbiased Risk는 Soft Threshold 방법을 통해 임계치 값을 선택하였다^[9]. 이 과정을 통해 심음의 잡음을 상대적으로 제거함으로써 주요한 정보만을 정제하여 심층학습의 입력이 될 수 있도록 전통적인 신호처리 기반의 전처리를 수행하였다.

Spectrogram은 소리 같은 신호의 주파수 스펙트럼 (Spectrum)을 시간과 주파수, 그리고 크기 정보 등을 축으로 시각화하는 방법이다^[8,10]. 일반적으로 2차원

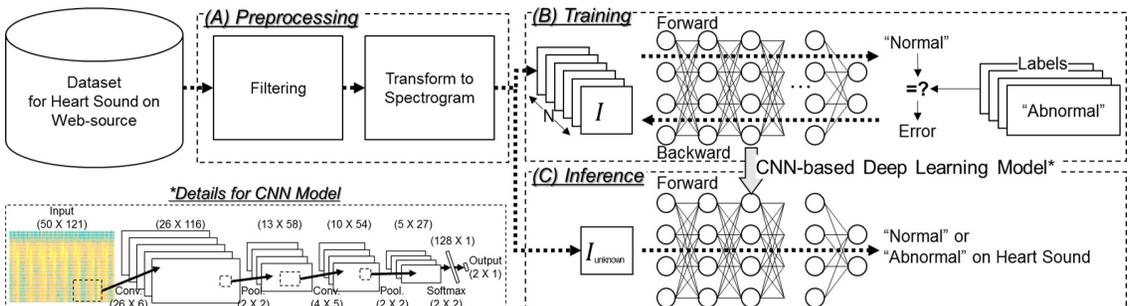


그림 3. 제안 방법의 개요
 Fig. 3. Overview of proposed method

입력을 학습하는 CNN 모델과의 연동을 위해 전처리된 청진음에 Spectrogram 변환 방법을 적용하여 CNN 모델에 입력하였다. 입력 처리의 편의성을 위해 가변적인 심음들을 10초 단위 나누어 고정된 크기의 입력으로 만들었으며, Window 크기 0.4초, Overlap 0.8초, 해상도 1024 변수를 통해 50×121 크기의 Spectrogram 기반의 2차원 입력으로 변환하였다. (그림 4)은 주어진 정상 혹은 비정상 심음에 대해 전처리 후, Spectrogram을 적용 변환한 결과의 예를 보인다.

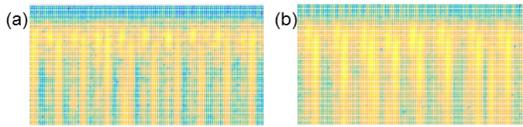


그림 4. 전처리된 Spectrogram 변환 예 (a)정상 (b)비정상
Fig. 4. Examples of preprocessed spectrogram (a)normal (b)abnormal

3.2 학습 (Training)

제안하는 심음 분석 진단 도움 알고리즘은 잘 알려진 CNN 모델에 기반으로 확장되었으며, 전처리된 2차원 심음을 입력으로 심층학습 기반의 모델을 학습하였다.

(그림 3)의 세부 CNN 모델에 나타나 있듯이 모델은 세부적으로 두 개의 Convolutional Layer^[6]와 Pooling Layer^[6]로 구성되었으며, 최종적으로 Fully Connected Layer^[6]와 Softmax Layer^[6]를 거쳐 심음의 정상 혹은 비정상, 두개의 결과를 출력한다. 주어진 학습 데이터 집합 (Training Dataset)을 이용하여 (수식 2)와 같이 실제 값과 모델에 의한 예측 값 간의 평균 제곱 오차로 정의된 손실함수 (Loss Function)

를 줄이는 방향으로 주어진 모델의 세부 학습 파라미터들을 조정함으로써 모델을 학습한다^[4,6].

$$E = \frac{1}{2} \sum_{i=1}^N (T_i - O_i)^2 \quad (2)$$

3.3 추정 (Inference)

학습된 CNN 모델을 이용하여 심음 측정 데이터 집합 (Test Dataset)의 정상 혹은 비정상 여부를 기준의 분류 문제의 성능을 판단하는데 빈번하게 사용되는 Accuracy, Recall, Precision, F1-score 측정 기준들을 이용함으로써 정량화하였다^[11]. 구체적인 구현 사항은 실험적 결과 부분을 통해 확인할 수 있다.

IV. 결 과 (Result)

우리는 공개된 PhysioNet/CinC Challenge 2016^[12]의 3126개 심음 데이터에 제안 방법을 적용하였다. 사용된 심음 데이터는 5~100초 동안 측정되었으며, 정상 혹은 비정상으로 구분된다. CNN 모델의 학습에 요구되는 충분한 Training Data를 위해 1~5% 사이의 임의의 Time Stretching, SNR 20 이상의 잡음 임의 삽입, 1~10% 사이의 Pitch Shifting 등의 잘 알려진 Data Augmentation 방법들을 적용하였다^[13].

실제 실험에서 사용된 CNN 모델의 세부 학습 사항은 Batch: 100, Optimizer: Adagrad, Learning Rate: 0.01, Epoch: 200이며^[4,6], 추정은 주어진 측정 데이터 집합에 교차검증 (10-fold Cross Validation)^[14]을 수행하였다. 모델은 Xeon 서버의 Ubuntu 환경에서 Torch^[15]를 통해 구현되었다.

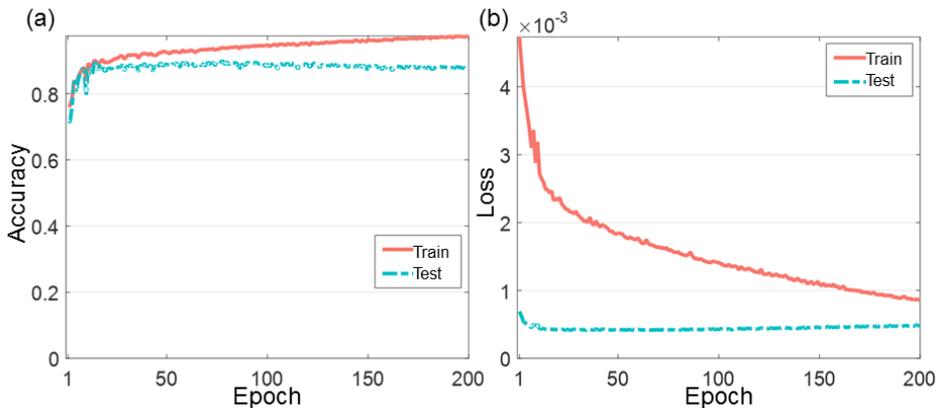


그림 5. 모델의 학습 및 검증 수행 결과
Fig. 5. Results of train and test on proposed method

위에 언급된 환경과 조건을 통해 구현된 실험 모델의 학습과 추정 진행 결과는 (그림 5)에 나타나 있으며, Accuracy 결과와 Loss 결과가 Epoch가 점차 증가함에 따라 모델 학습이 수렴함에 따라 적절한 모델의 학습과 추론이 수행되었음을 확인할 수 있다.

(그림 6)의 제안 방법의 결과는 Accuracy, Recall, Precision, F1-score 4가지의 측정 지표를 이용하여 10-fold Cross Validation의 평균값으로 나타내었다. 또한, 신호처리 기반의 잡음 제거와 같은 전처리의 타당성을 확인하기 위해 동일한 CNN 모델에 대해 처리 유무에 대한 성능 비교 실험도 수행하였다.

실험 결과, 전반적인 측정 지표에서 전처리의 신호 처리 후 모델에 적용한 결과가 다소 높은 성능 지표를 가짐을 통해 제안 방법의 전처리 타당성을 확인할 수 있었고, 4가지 지표 모두 0.8이상의 결과는 제안하는 방법이 주어진 심음 이상 유무 분류에 도움이 될 수 있음을 확인할 수 있었다.

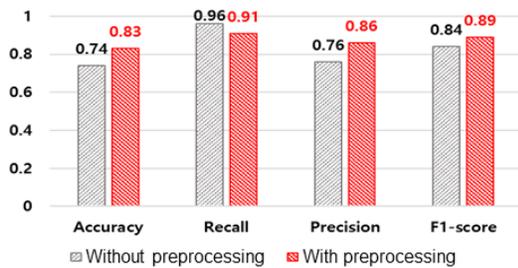


그림 6. 심음 이상 여부 판단 성능 측정 결과
Fig. 6. Performance of abnormality classification on cardiac sound dataset

V. 결 론 (Conclusion)

우리는 공개된 심음 데이터 집합에 대해 제안 방법을 구현하여 80% 이상의 성능 지표를 가짐을 확인함으로써 제안 방법이 심음 이상 유무 판단에 도움을 줄 수 있음을 확신하게 되었다. 하지만 사용된 공개 심음 데이터 집합이 상대적으로 잡음이 적은 이상적인 환경에서 측정되었다는 점 등의 실제 청진이 되는 환경과의 차이가 실제 환경 적용에 있어서 성능 열화 요인으로 예상된다.

향후, 실제 모바일 전자 청진기와와의 연동을 통해서 잡음이 있는 환경에서 측정된 심음 데이터 집합을 추가적으로 학습하여 일상적인 상황에서 사용될 수 있도록 Generative Deep Learning 모델을 고려한 연구들도 요구된다. 또한, 본 연구에서 적용된 모델을 심

음 외의 다양한 생체신호를 분석함으로써 스마트폰과 같은 모바일 환경에서의 포괄적이고 범용적인 진단 도움 기술의 가능성을 고려해 볼 수 있을 것으로 예상된다.

결과적으로 우리가 제안한 방법은 심층학습 기반의 심장 순환기 생체 신호 분석 진단 도구에 대한 가능성을 확인할 뿐만 아니라 심음 외의 EEG, PPG와 같은 다양한 생체 신호 분석에 대한 심층학습을 이용한 분석의 가능성도 가늠해볼 수 있다.

References

- [1] J.-H. Kim, S.-M. Lee, and K.-H. Park, "Stepwise detection of the QRS complex in the ECG signal," *J. KICS*, vol. 41, no. 2, pp. 244-253, 2016.
- [2] M. L. Rice and D. J. Doyle, "Comparison of phonocardiographic monitoring locations," *IEEE 17th Annu. Conf. Eng. in Med. and Biology Soc.*, vol. 1, pp. 685-686, Sept. 1995.
- [3] K.-Y. Park, S. Hong, J. Lee, J. Park, and E. Jung, "Animal diagnosis system using wireless digital stethoscope," *J. KICS*, vol. 38, no. 9, pp. 722-727, 2013.
- [4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*, MIT Press, 2016.
- [5] H. Han and K.-Y. Song, "Electroencephalogram-based driver drowsiness detection system using errors-in-variables(EIV) and multilayer perceptron (MLP)," *J. KICS*, vol. 39, no. 10, pp. 887-895, 2014.
- [6] Y. LeCun, *LeNet-5, convolutional neural networks*(2015), Retrieved June, 1, 2016 from <http://yann.lecun.com/exdb/lenet>.
- [7] K. Lee, Y. Ji, Y. Jeon, and Y. C. Park, "Development and implementation of noise-canceling technology for digital stethoscope," *J. Biomedical Eng. Res.*, vol. 34, no. 4, pp. 204-211, 2013.
- [8] R. Cohen, *Signal denoising using wavelets*, Project Report, Department of Electrical Engineering Technion, Israel Institute of Technology, Haifa, 2012.
- [9] F. Luisier, T. Blu, and M. Unser, "A new SURE approach to image denoising: Interscale

orthonormal wavelet thresholding,” *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 593-606, 2007.

- [10] L. Deng, O. Abdel-Hamid, and D. Yu, “A deep convolutional neural network using heterogeneous pooling for trading acoustic invariance with phonetic confusion,” *2013 IEEE Int. Conf. Acoustics, Speech and Sign. Process.*, pp. 6669-6673, Vancouver, BC, Canada, May 2013.
- [11] D. M. Powers, “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation,” *J. Machine Learning Technol.*, vol. 2, no. 1, pp. 37-63, 2011.
- [12] PhysioNet/CinC Challenge 2016. Challenge Data(2016), Retrieved June 1, 2016, from <https://www.physionet.org/challenge/2016>.
- [13] J. Schlüter and T. Grill, “Exploring data augmentation for improved singing voice detection with neural networks,” *ISMIR*, pp. 121-126, Malaga, Spain, Oct. 2015.
- [14] C. M. Bishop, *Pattern recognition and machine learning*, Springer, 2006.
- [15] R. Collobert, S. Bengio, and J. Mariéthoz, *Torch: a modular machine learning software library*, No. EPFL-REPORT-82802, Idiap, 2002.

이 재 구 (Jaekoo Lee)



2009년 2월 : 숭실대학교 정보통신전자공학부 졸업
 2011년 4월 : UC San Diego, Electrical and Computer Engineering 석사
 2011년~2013년 LG전자 주임 연구원

2013년 9월~현재 : 서울대학교 전기정보공학부 박사과정
 <관심분야> 기계학습, 딥러닝, 그래프분석, 시계열 분석

김 현 재 (Hyunjae Kim)



2016년 2월 : 서울대학교 전기정보공학부 졸업
 2016년 3월~현재 : 서울대학교 전기정보공학부 석박통합과정
 <관심분야> 딥러닝, 시계열 분석

김 성 원 (Sungwon Kim)



2016년 8월 : 서울대학교 전기정보공학부 졸업
 2017년 3월~현재 : 서울대학교 전기정보공학부 석박통합과정
 <관심분야> 딥러닝

송 종 윤 (Jongyoon Song)



2017년 2월 : 서울대학교 전기정보공학부 졸업
 2017년 3월~현재 : 서울대학교 전기정보공학부 석박통합과정
 <관심분야> 딥러닝, 자연어 처리

윤 성 로 (Sungroh Yoon)



1996년 2월 : 서울대학교 전기
공학부 졸업

2002년 : Stanford University,
Electrical Engineering 석사

2006년 : Stanford University,
Electrical Engineering 박사

2006년 : Stanford University,
박사후 연구원

2006년~2007년 : 미국 인텔 선임연구원

2007년~2012년 고려대학교 전기전파공학부 조교수

2012년 9월~현재 : 서울대학교 전기정보공학부 부교
수/조교수

<관심분야> 딥러닝, 빅데이터 기반 인공지능, 생물
정보학, 지능형 임베디드시스템 등