

합성곱-장단기 기억 신경망의 하이브리드 결합 모델을 이용한 부정맥 분류

조익성¹ · 권혁승^{2*}

Arrhythmia Classification using Hybrid Combination Model of CNN-LSTM

Ik-Sung Cho¹ · Hyeog-Soong Kwon^{2*}

¹Assistant Professor, School of Interdisciplinary Studies, Daegu University, Daegu, 38453 Korea

^{2*}Professor, Department of IT Engineering, Pusan National University, Miryang, 50463 Korea

요약

부정맥은 심장 박동이 비정상 혹은 불규칙하게 뛰고 있는 상태를 말하며, 실신이나 심장돌연사 등과 같은 위험한 상황을 유발할 수 있기 때문에 이의 조기 검출은 매우 중요하다.

하지만 심전도 신호의 개인차로 인해 분류 시 성능하락이 나타날 수밖에 없다. 본 연구에서는 CNN-LSTM 하이브리드 결합 모델을 이용한 부정맥 분류 방법을 제안한다. 이를 위해 먼저 잡음을 제거한 ECG 신호에서 R파를 검출하고 단일 비트 세그먼트를 추출하였다. 이후 부정맥 신호의 특징을 세밀하게 추출하도록 8개의 합성곱 계층으로 구성하고 이를 LSTM의 입력으로 사용한 후 가중치를 학습시키고 검증 데이터로 모델을 평가한 후 정상 및 부정맥 분류의 변화를 확인하였다. 제안한 방법의 타당성 검증을 위해 MIT-BIH 부정맥 데이터베이스를 사용하여 정확도(accuracy), 정밀도(precision), 재현율(recall), F1 스코어가 사용되었다. 성능평가 결과, 정확도, 정밀도, 재현율, F1 스코어는 각각 92.3%, 90.98%, 92.20%, 90.72%의 우수한 분류율을 나타내었다.

ABSTRACT

Arrhythmia is a condition in which the heart beats abnormally or irregularly, early detection is very important because it can cause dangerous situations such as fainting or sudden cardiac death. However, performance degradation occurs due to personalized differences in ECG signals.

In this paper, we propose arrhythmia classification using hybrid combination model of CNN-LSTM. For this purpose, the R wave is detected from noise removed signal and a single bit segment was extracted. It consisted of eight convolutional layers to extract the features of the arrhythmia in detail, used them as the input of the LSTM. The weights were learned through deep learning and the model was evaluated by the verification data.

The performance was compared in terms of the accuracy, precision, recall, F1 score through MIT-BIH arrhythmia database. The achieved scores indicate 92.3%, 90.98%, 92.20%, 90.72% in terms of the accuracy, precision, recall, F1 score, respectively.

키워드 : 부정맥, 합성곱 신경망, 장단기 기억 신경망, 심층 신경망, MIT-BIH

Keywords : Arrhythmia, CNN, LSTM, Deep learning, MIT-BIH

Received 10 October 2021, Revised 14 October 2021, Accepted 26 October 2021

* Corresponding Author Hyeog-Soong Kwon(E-mail:hskwon@pusan.ac.kr Tel:+82-55-350-5411)

Professor, Department of IT Engineering, Pusan National University, Miryang, 50463 Korea

Open Access <http://doi.org/10.6109/jkiice.2022.26.1.76>

print ISSN: 2234-4772 online ISSN: 2288-4165

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

부정맥은 심장 박동이 비정상적으로 빨라지거나 늦어지거나 혹은 불규칙하게 뛰고 있는 상태를 말한다[1]. 정상 맥박수는 분당 60-100회로, 맥박수가 분당 50회 미만이면 '서맥', 분당 100회 이상이면 '빈맥'이라고 부른다. 부정맥은 증상이 즉시 나타나는 경우도 있지만, 간헐적으로 나타난 후 저절로 소실되는 경우가 많기 때문에 적시에 발견해 치료하지 않으면 뇌졸중, 심정지 등을 유발하여 사망에 이를 수 있다. 따라서 조기 진단을 통한 관리가 무엇보다 중요하다[2].

기존 연구에서는 웨이블릿 변환 (wavelet transform), 통계적(statistical) 방법, 형태학적(morphological) 방법 등의 특징점을 이용한 부정맥 분류 방법이 제안되었다 [3, 4]. 하지만, 특징 추출 과정에서 잡음(signal noise)이 포함될 경우 분류시 성능 하락이 나타날 수 있다.

최근에는 컴퓨터 성능의 발전으로 회귀(regression) 및 분류 문제를 해결하는 데 있어 머신러닝(machine learning), 합성곱 신경망(CNN : Convolutional Neural Network), 순환 신경망(RNN : recurrent neural network) 등의 딥러닝(deep learning)을 이용한 부정맥 분류 방법이 제안되었다. 하지만 합성곱 신경망은 시간적 특성을 반영하기 힘들다는 단점이 있으며 순환신경망은 학습이 진행될수록 학습능력이 현저히 줄어드는 장기 의존성 문제가 발생하게 된다. 장단기 기억 신경망(LSTM : Long Short-Term Memory)은 기존의 RNN 모델에서 장기간(long-term) 학습 시 생기는 기울기(gradient)의 소실과 폭발 문제를 개선하기 위해 제안되었다[5-8].

본 연구에서는 CNN-LSTM 하이브리드 결합 모델을 이용한 부정맥 분류 방법을 제안한다.

이는 심실의 탈분극 상태를 반영하는 QRS 콤플렉스와 같이 심전도 신호의 지역적인 특징을 추출하는데 효과적인 CNN과 순환신경망의 장기 의존성 문제를 해결할 수 있는 LSTM을 연결하여 이미지와 시계열 데이터를 동시에 고려한 분류방법이다.

이를 위해 먼저 잡음을 제거한 ECG 신호에서 R파를 검출하고 단일 비트 세그먼트를 추출하였다. 이후 부정맥 신호의 특징을 세밀하게 추출하도록 8개의 합성곱 계층으로 구성하고 이를 LSTM의 입력으로 사용한 후 가중치를 학습시키고 검증 데이터로 모델을 평가한 후 정상 및 부정맥 분류의 변화를 확인하였다. 제안한 방법

의 타당성 검증을 위해 MIT-BIH 부정맥 데이터베이스를 사용하여 정확도(accuracy), 정밀도(precision), 재현율(recall), F1 score가 사용되었다.

본 논문의 구성은 다음과 같다. 2장에서는 딥러닝 모델, 3장에서는 본 연구의 제안 모델에 대하여 4장에서는 실험결과에 대하여 논하고 5장에서 결론으로 마무리한다.

II. 딥러닝 모델

2.1. 합성곱 신경망

합성곱 신경망(CNN : Convolutional Neural Network)은 데이터에서 핵심 특징들을 추출하여 분류 및 예측하는 딥러닝 모델이다. 주로 이미지 분야 연구에 사용되어 왔으며, 최근에는 의료영상, 생체신호 분석 연구에 폭넓게 사용되고 있다.

CNN은 심층신경망에 합성곱 계층과 풀링 계층이 결합된 구조로 일반적인 구조는 그림 1과 같다.

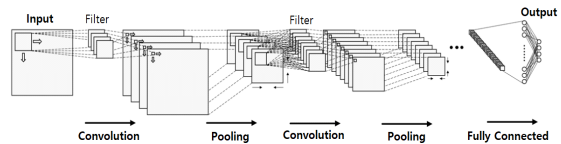


Fig. 1 CNN structure

합성곱, 풀링 계층을 거쳐 완전연결계층(fully connected layer)을 통해 최종 결과를 출력하고 클래스 분류 시 소프트맥스(softmax) 함수를 통해 확률 값으로 변환한다.

일반적으로 신경망의 층이 깊어질수록 더욱 많은 정보를 학습할 수 있지만 학습 데이터의 지역적 특성에만 적용하는 과다학습(overfitting) 문제가 발생하게 된다. 이를 해결하기 위해 뉴런을 임의로 삭제하면서 학습하는 드롭아웃(dropout) 방법을 사용한다. 이는 학습시 은닉층의 뉴런을 무작위로 골라 삭제하고 삭제된 뉴런은 신호를 전달하지 않게 하여 일부 특징이 특정 뉴런들에 고정되는 것을 막아 가중치의 균형을 잡는다.

하지만 CNN의 학습과정은 순서가 중요하지 않은 정보들이 공통으로 가지는 특징들만 관심이 있기 때문에 실시간으로 들어오는 정보들의 순서 관계를 처리할 수 없는 문제점이 있다. 이에 과거 및 현재 정보를 기반으로 미래 정보를 예측하기 위해 시계열 특징을 반영한

RNN이 등장하였다[9].

2.2. 순환 신경망

순환신경망(RNN : Recurrent Neural Network)은 기존의 FFNN(Feed Forward Neural Networks)의 확장으로 변동이 심한 시계열 데이터를 처리하는데 적합한 모델이다. 구조적으로는 현재와 다음 시점을 연결하여 여러 개의 데이터가 순서대로 입력되었을 때 이전 시점의 데이터가 유지되도록 반복(recurrent)의 개념이 추가된 것으로 계열 데이터의 패턴과 특성을 파악하는데 적합한 모델이다. 이러한 특성 때문에 기존의 FFNN 모델의 학습을 위해 사용하였던 역전파(back propagation)와는 조금 다르게 시간에 걸쳐 풀어서 해석하는 BPTT(Back-Propagation Through Time) 방식을 사용한다.

RNN은 기존의 인공신경망 구조와는 다르게 모든 과정에 대해 파라미터(parameter)값을 공유하기 때문에, 각 단계(step)마다 입력(input)값만 다를 뿐 거의 똑같은 계산을 하고 있다는 것을 보여준다. 그렇기 때문에 학습해야 할 파라미터의 수를 많이 줄여준다.

하지만 기존의 RNN 모델은 이전 시점의 출력값이 현재 시점의 값에 영향을 주기 때문에, 입력값이 길어지게 되면 출력값들이 곱해지는 단계가 반복되므로 누적값이 증가하는 폭발(explode) 또는 0으로 수렴하는 소실(vanish) 문제가 발생하여 학습이 어려워진다[10].

2.3. 장단기 기억 신경망

장단기 기억 신경망(LSTM : Long Short-Term Memory) 모델은 기존의 RNN 모델에서 장기간(long-term) 학습 시 생기는 기울기(gradients)의 소실과 폭발 문제를 개선하기 위해 제안되었다.

RNN과 비슷한 구조를 가지고 있지만 세 개의 게이트(Forget, Input, Output)를 이용하여 정보를 제어하며 이를 셀 스테이트(cell-state)를 통해 각 게이트의 연산된 값을 전달하게 된다. 따라서 각 단계에서 문맥(content)을 덮어쓰는 기존의 RNN과는 달리 게이트를 통해 기존의 기억을 유지할 것인지, 아니면 삭제할 것인지를 결정하기 때문에 장, 단기간의 정보를 쉽게 전달할 수 있게 된다. 그림 2와 같이 LSTM에서 셀 스테이트는 핵심 역할로서 선형 연산을 거쳐 전체 라인을 통과한다. 이러한 구조로 인해 정보가 유지되면서 순차적으로 다음 단계로 전달된다.

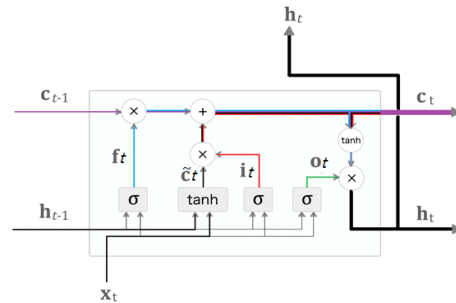


Fig. 2 Cell-state of LSTM

LSTM은 세 개의 게이트로 구성됨으로 순환신경망 알고리즘과는 달리 LSTM은 편도함수의 곱이 입력되는 노드 수만큼 계산되는 것을 방지함으로써 보다 정확한 가중치 값을 획득하게 된다.

LSTM의 가장 첫 번째 과정인 망각(forget) 게이트는 과거의 정보를 얼마나 잊을지 또는 기억할지를 결정하는 단계이다.

수식(1)과 같이 망각 게이트(f_t)에서는 이전 시점의 출력값(h_{t-1})과 현재 시점의 입력값(x_t)에 가중치(W_f)를 곱하여 sigmoid(σ) 활성화 함수를 이용해 0과 1사이의 값으로 출력한다. 이는 0에 가까울수록 많은 정보를 잊은 것이고 1에 가까울수록 많은 정보를 기억하는 것이며 중간에 있는 값은 가능한 모든 후보를 나타낸다.

계산된 f_t 는 이전 기억셀(C_{t-1})과 원소별 곱을 취하게 된다. 여기서 t 는 시간표기(time stamp)를 의미하며, W 는 가중치(weight), b 는 편향(bias)를 의미한다.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

망각 게이트가 과거의 정보를 잊을지 또는 기억할지를 결정하는 과정이었다면 입력 게이트는 새로운 정보를 셀 스테이트에 저장될지를 결정하는 과정으로 두 단계로 이루어져 있다.

수식 (2)와 같이 i_t 는 이전 시점의 출력값(h_{t-1})과 현재 시점의 입력값(x_t)을 통해 시그모이드 연산을 진행하는 단계와 새로운 기억 셀 스테이트(\tilde{C}_t)에 정보를 추가하기 위해 이전 시점의 출력값(h_{t-1})과 현재 시점의 입력값(x_t)에 하이퍼볼릭 탄젠트(hyperbolic tangent) 연산, 즉, -1과 1 사이의 값을 가지는 함수로 취하여 새로 추가되는 각 원소가 정보로써 어느 정도의 가치가 있는지를 판단하는 단계를 거친다.

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \end{aligned} \quad (2)$$

이후 망각 게이트와 입력 게이트를 통해 얻어진 값들을 바탕으로 새로운 스테이트로 갱신(update)하여 C_t 를 생성하는 단계를 거친다.

수식 (3)에서 C_{t-1} 은 이전 셀을 의미하는데 여기에 f_t (forget gate)에서 출력된 값을 원소별 곱하게 된다. 여기서 \times 는 행렬곱이 아니라 원소별 곱을 의미한다. 특이한 점은 각각의 다른 가중치와 sigmoid, tanh를 활성화 함수로 \tilde{C}_t 와 i_t 를 연산하여 기존의 값에 영향을 주게 된다.

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (3)$$

이렇게 출력된 두 가지 값을 이용해 다음 셀 스테이트에 얼마나 반영시킬지 결정하게 된다.

최종적으로 수식 (4)와 같이 출력 게이트와 수식 (3)의 새로운 셀 스테이트값 C_t 를 곱하고 다음 스테이트에 어느 정도 반영할지를 계산하여 최종 출력값(h_t)를 생성한다.

$$\begin{aligned} o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t \times \tanh(C_t) \end{aligned} \quad (4)$$

결론적으로 LSTM은 기존의 RNN과 달리 기억 셀이라는 것이 존재하고 이를 활용하기 위해 3개의 게이트가 사용된다. 이 기억셀은 역전파시 덧셈과 곱셈 노드만을 지나고, 덧셈 노드는 역전파시 전해지는 기울기를 그대로 흘리기 때문에 기울기 소실이 발생하지 않으며

역셀에서 사용되는 곱셈 노드는 행렬 곱이 아닌 원소별 곱이기 때문에 매번 새로운 게이트 값을 사용하므로 곱셈 효과가 누적되지 않아 기울기 소실이 발생하지 않는다[11-12].

III. 제안 모델

본 연구에서 제안하는 모델은 신호의 형태학적 특징을 추출하는데 적합한 1차원 합성곱 신경망(CNN)과 시간적 맥락을 기억하여 기존 순환신경망의 장기 의존성 문제를 개선하기 위한 LSTM 구조를 결합한 신경망이다. 이는 심전도 신호의 공간적, 시간적 특성 정보를 작은 반복학습으로 효과적으로 학습하여 최적의 분류율을 나타내기 위한 방법이다.

3.1. 전처리 및 분할

본 연구에서는 피지오넷(physionet)에서 제공하는 MIT-BIH 부정맥 데이터베이스를 사용하였다. 각 레코드에 해당하는 클래스는 AAMI(Association for the Advancement of Medical Instrumentation)이 권고에 따라 페이스메이커(pacemaker) 비트가 포함되어 있는 4개의 레코드는 제외하고 N, S, V, F, Q의 5개로 분류하였다. 그림 3과 같이 N 클래스는 정상 심장 박동에 해당하는 그룹, S 클래스는 심방(atrium) 또는 방실결절(atrioventricular node)에 위치한 심실위(supraventricular) 이소성 박동에 해당하는 그룹, V 클래스는 심실(ventricular)과 관련된 이소성 박동 종류를 모아놓은 그

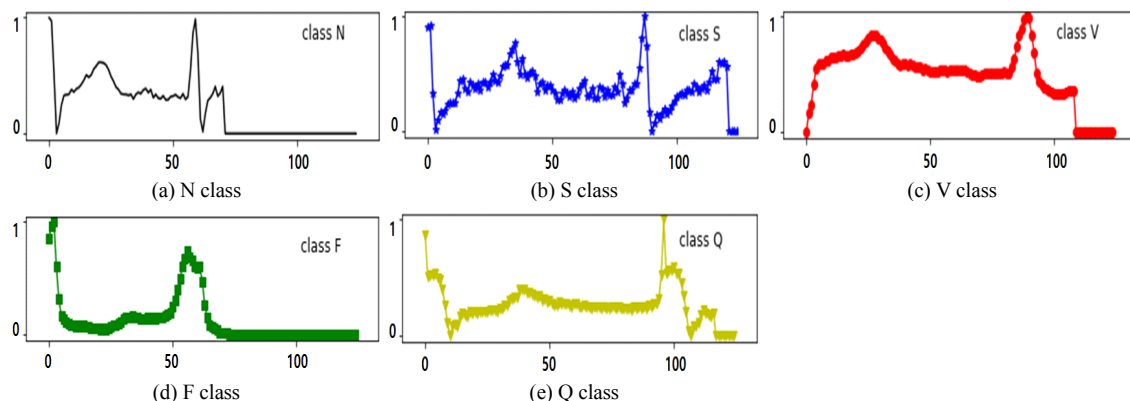


Fig. 3 (a) N class, (b) S class, (c) V class, (d) F class (e) Q class

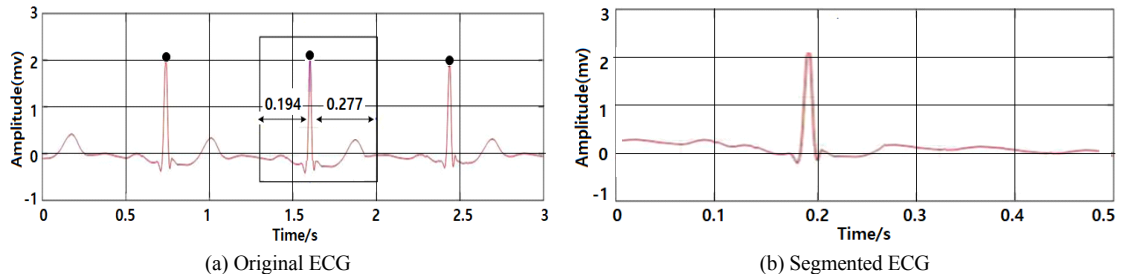


Fig. 4 Split structure of single beat

를, F 클래스는 이소성 박동이 심실에 도달하는 정상적인 전도와 일치하여 정상과 이소성 박동의 형태가 복잡적(complex)으로 나타나는 융합(fusion) 박동 종류를 모은 그룹을 의미한다. Q 클래스는 인공 심박동기인 페이스메이커(pacemaker)의 전기자극(electrical stimulation)에 의해 심장박동이 제어(control)되는 상태를 일컫는 페이스(paced) 박동(beat), 분류 불가능한(unclassifiable) 박동, 페이스 박동과 정상 박동이 융합된 박동(fusion of paced and normal beat)과 같이 클래스를 알 수 없는(unknown beat) 박동에 해당하는 심장박동 종류를 모아 놓은 그룹이다.

48개의 레코드 중 페이스메이커에 의한 비트가 포함되어 있는 4개의 레코드는 제거되었으며 44개의 레코드는 각각 학습 및 검증 데이터에 할당되었다. 전체 레코드 중 22개의 레코드에는 일반적으로 정상적인 Normal 비트가 많이 포함된 경향이 있으며, 나머지 22개의 레코드는 복잡한 심실, 접합부, 상심실성 부정맥 등의 정보가 포함되어 있다. 따라서 성능 테스트의 신빙성을 높이기 위해 학습 및 검증 데이터에 분포한 클래스 비율을 맞추어 레코드 셋을 22개씩 각각 할당하였다. 이 데이터 셋은 실험에 활용하기에 적합하도록 모든 심전도 데이터를 비트별로 잘라내고 제로패딩 과정을 수행하여 모든 비트의 길이를 동일하게 맞춰놓았다.

제안 모델은 1D CNN을 통해 심전도의 공간적 특징을 추출하고 이를 시계열 데이터의 순차적 특징을 고려할 수 있는 LSTM의 입력 자료로 사용한다. 즉 합성곱 연산에 사용될 ECG 신호의 형태학적 특징을 포함해야 한다. 이를 위해 그림 4와 같이 R 피크 위치 전 0.194, 후 0.277초로 단일 비트로 분할하여 QRS영역을 포함한 심전도의 특징 파형을 포함하도록 구성하였다.

이후 한 비트씩 슬라이딩(sliding)하며 단일 비트를 세 비트로 묶어 1.38초씩 구성하였다. 1.38초 내에 포함

된 세 개의 비트가 모두 N 클래스일 경우, 정상, 그렇지 않으면 부정맥으로 구성 하였다. 또한, 학습 및 검증을 위한 데이터 셋으로 MIT-BIH 부정맥 데이터베이스에서 정상과 부정맥을 별도로 각각 44,085, 10,890비트로 전체 데이터 54,975비트 샘플을 학습과 검증을 위해 8대 2로 분리하여 제안 모델의 최종 성능을 평가하였다.

학습 및 검증을 위한 데이터 클래스 분포는 표 1과 같다.

Table. 1 Training and validation data set

Class \ Data	Training samples	Validation samples	Total samples
Normal	35,268	8,817	44,085
Arrhythmia	8,712	2,178	10,890
Total	43,980	10,995	54,975

3.2. CNN-LSTM 딥러닝 모델

제안한 CNN-LSTM 하이브리드 결합 모델의 전체 흐름도는 그림 5와 같다. 본 연구에서는 딥러닝 기반 알고리즘인 CNN과 LSTM을 직렬 병합하여 특정 구간에서의 부정맥을 예측하고자 하였다.

제안 모델에 공급되는 입력값은 앞서 언급한 세 비트 신호와 재구성된 이진(binary) 클래스이다. 심전도 데이터는 1차원 데이터로 구성되어 있으므로 특징 벡터(vector)값을 추출하기 위하여 1차원 합성곱 계층(1D convolution layer)을 사용하였다.

제안 모델은 1D CNN을 통해 심전도의 공간적 특징을 추출하고 이를 시계열 데이터의 순차적 특징을 고려할 수 있는 LSTM의 입력 자료로 이용함으로써 시공간 특징을 모두 분류 과정에 사용할 수 있다. 즉 합성곱 연산으로 추출된 1차원의 공간 특징을 이용한다는 점에서 원 신호를 1차원화하여 입력 자료로 이용하는 기존 LSTM과 차별성을 갖는다.

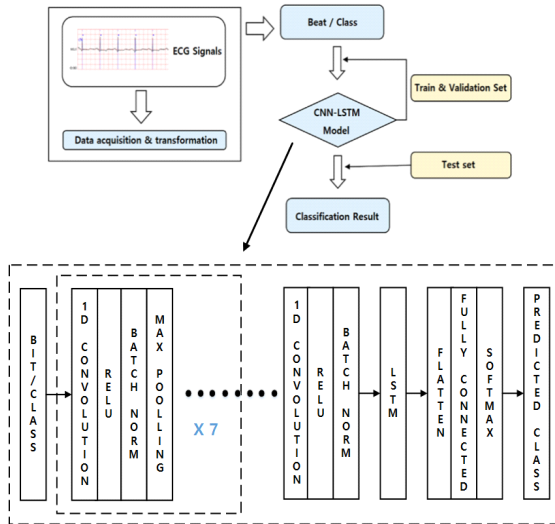


Fig. 5 Proposed CNN-LSTM model

합성곱 계층은 합성곱 연산을 수행하여 특징 맵(map)을 추출하는 계층이다.

합성곱을 수행할 때의 윈도우 크기(window size)를 의미하는 커널 크기(kernel size) 값과, 합성곱 과정에서 이동하는 보폭 값을 정하는 스트라이드(stride) 값을 이용하여 신호의 처음부터 마지막까지 해당 ECG 신호의 특징을 가진 필터를 생성한다. 본 연구에서는 8개의 합성곱 계층과 커널 크기는 2로 설정하였다. 스트라이드 값은 합성곱 연산을 할 때 1만큼 이동할 수 있도록 1로 설정하였다. 각 계층에는 ReLU 활성화 함수를 사용하였다.

또한 7번째 합성곱 계층까지는 최대 풀링(max pooling)을 사용하여 입력 값의 길이를 효과적으로 줄임과 동시에 주요 특징들을 추출하도록 하였다. 마지막 합성곱 계층은 LSTM으로 들어갈 충분한 정보를 위해 최대 풀링은 두지 않았다.

제안 모델에서 사용된 LSTM의 출력 은닉 벡터(hidden vector)는 64로 생성되었다. 이는 CNN으로부터 추출된 특징들 사이의 시간적 관계(temporal relationship)를 인코딩(encoding)한 값이다. 이후 이 값을 플래튼 계층(flatten layer)을 통해 1차원으로 변환한 후, 전결합 계층(fully connected layer)으로 전달되도록 한다. 전결합 계층의 최종 출력 값은 소프트맥스 함수를 통해 정상 및 부정맥에 대하여 해당하는 가장 높은 확률 값이 선택된다. 출력값의 합은 항상 1이며, 가장 큰 출력 값에 해당하는 클래스를 실제 정답과 비교하여 손실(loss)값을 갱

신(update)한다.

3.3. 모델의 학습

제안 모델의 파라미터는 matlab과 python의 반복적인 실험을 통해 경험적으로 선정된 값들을 사용하였으며 출력할 합성곱 필터의 파라미터는 표 2와 같다.

Table. 2 Convolution filter parameter

Convolution Filter	Value
1st convolution filter	16
2nd convolution filter	16
3rd convolution filter	32
4th convolution filter	32
5th convolution filter	64
6th convolution filter	64
7th convolution filter	128
8th convolution filter	256

손실 결과를 통해 가중치를 갱신(update)할 최적화 함수는 안정적으로 학습할 수 있는 Adam을 사용하였으며 0.001의 학습률이 적용되었다. 학습률이 너무 클 경우 최저치에 수렴하지 않을 수 있으며, 너무 작을 경우 많은 학습시간이 소요되어 적절한 값을 선택해야 한다. 학습을 위한 손실 함수는 이진 분류에 사용되는 이진 교차 엔트로피(binary cross-entropy, BCE)가 적용되었다.

학습을 위한 배치 크기(batch size)는 64로 설정하였다. 배치 크기란 몇 개의 샘플만큼 클래스를 예측하고 실제 클래스와 비교 할지 결정하는 하이퍼 파라미터로, 너무 작을 경우 가중치 갱신이 자주 일어나 불안정하게 학습될 수 있으며, 너무 클 경우 처리해야 양이 많아 학습 속도가 느려질 수 있다.

전체 데이터에 대한 학습 횟수를 의미하는 epoch는 1 이상부터는 성능 하락이 나타났기 때문에 그림 6과 같이 최고 성능을 나타내는 10으로 설정하였다.

IV. 실험결과 및 고찰

4.1. 모델 검증

CNN-LSTM 모델은 크게 학습 과정(training)과 검증 과정(validation)으로 나눌 수 있다. 학습 과정은 매개변수 갱신이 목표이며 검증 과정은 갱신된 매개변수를 통

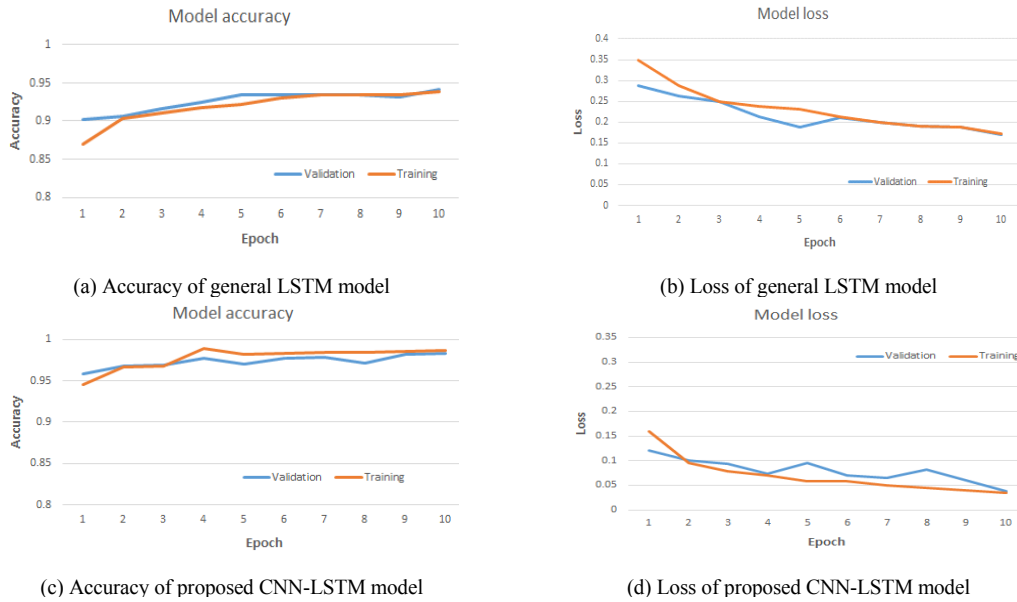


Fig. 6 Accuracy and loss of general LSTM & proposed CNN-LSTM model

해 성능을 확인하는 과정이다. 본 연구에서는 학습 데이터와 검증 데이터의 비율을 8:2으로 설정하였다. 매개변수는 1 epoch 마다 한번 갱신되며 검증 데이터 셋을 통해 그 정확도를 확인하였다.

그림 6은 MIT-BIH 부정맥 데이터베이스를 대상으로 epoch에 따른 일반 LSTM과 제안한 CNN-LSTM 모델의 분류 정확도(accuracy)와 손실(loss)을 비교한 것이다. 본 연구에서 제안하는 하이브리드 결합 모델은 작은 반복학습으로 효과적으로 최적의 분류율을 나타내기 위한 방법이다.

학습 및 검증 데이터의 각 모델 결과를 분석한 결과, 대부분의 레코드는 10 epoch일 때 학습과 검증에 대한 정확도(accuracy)는 각각 0.987, 0.983, 학습과 검증에 대한 손실(loss)은 0.035, 0.038로 각각 수렴하고 있는 것으로 확인되었다. 그 이상의 경우 분류 정확도 변화가 일어나지 않았다.

QRS 영역은 심전도의 지역적 특징이 매우 뚜렷하기 때문에 이러한 공간적 특징을 추출할 수 있는 1D CNN을 결합한 하이브리드 모델이 더 나은 성능을 나타내는 것으로 판단된다.

4.2. 분류 성능 평가

본 연구에서 제안한 CNN-LSTM 모델의 최종 성능

평가를 위해 시험 데이터를 통해 분류 성능을 평가하였다. 분류를 위한 성능 평가는 일반적으로 양성 참(TP), 음성 참(TN), 양성 거짓(FP), 음성 거짓(NP)의 4가지 측정 요소로 구한다.

성능 지표로는 모든 분류 결과 중에서 올바르게 분류된 비율을 나타내는 정확도(accuracy, acc), 양성으로 분류한 것 중에서 실제 양성 비율을 나타내는 정밀도(precision), 모든 검출된 사건 중에서 올바르게 양성으로 분류한 사건의 비율을 나타내는 재현율(recall, sen), 재현율과 정밀도의 조화 평균을 나타내는 F1 점수가 사용되었다. 성능 지표는 아래 수식과 같다.

$$Accuracy(\text{정확도}) = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

$$Precision(\text{정밀도}) = \frac{TP}{TP + FP}$$

$$Recall(\text{재현율}) = \frac{TP}{TP + FN}$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

본 연구는 정상 신호로 분류되었는지, 이상 신호, 즉 부정맥으로 분류되었는지를 확인하기 위해 클래스 수로 평균을 취하는 매크로(macro) 평균을 사용하였다. 표 3은 각 모델에 따른 성능 평가 결과를 나타낸다.

성능 평가 결과 제안하는 모델의 정확도(accuracy)와 정밀도(precision)는 각각 92.3%, 90.98%로 동일한 환경에서 실험한 일반 LSTM 모델에 비해 성능이 우수한 것을 확인할 수 있다.

표에서 볼 수 있듯이 타 모델 대비 본 연구에서 제안한 CNN-LSTM 하이브리드 결합 모델의 성능이 가장 우수한 것을 확인하였다.

Table. 3 Result of performance test

	accuracy	precision	recall	F1
General LSTM	90.57%	89.20%	88.74%	88.81%
Proposed CNN-LSTM	92.3%	90.98%	92.20%	90.72%

이는 앞서 언급한 바와 같이 CNN에서의 QRS 형태 변화에 대한 공간적 특성을 더 국소적으로 포착할 수 있는 특성이 LSTM의 학습에 효율적으로 반영된 결과라고 할 수 있다.

표 4는 기존 부정맥 분류 모델 및 본 연구의 방법 비교 및 정확도 성능 기준 결과를 나타낸다. 본 연구에서 제안한 방법은 심전도 신호의 지역적인 특징을 추출하는데 효과적인 CNN과 순환신경망의 장기 의존성 문제를 해결할 수 있는 LSTM을 연결하여 이미지와 시계열 데이터를 동시에 고려한 분류방법이다.

Singh은 LSTM 기반 모델로 부정맥을 분류하였으며, Sellami은 1D CNN 구조와 배치 가중치 손실을 적용하여 부정맥을 분류하였다[13-14].

본 연구는 Sellami의 연구에서 사용된 ECG 단일 beat 분할 방법을 적용하여 전처리하고, Singh이 제안한 분류 방법을 결합하여 심전도 신호의 공간적 특징을 추출하고 이를 시계열 데이터의 순차적 특징을 고려할 수 있는 LSTM의 입력 자료로 이용함으로써 시공간 특징을 모두 분류 과정에 사용할 수 있는 하이브리드 모델을 제안하고, 일반 LSTM 모델과의 성능 비교를 통해 제안하는 모델이 우수함을 확인하였다.

Table. 4 Comparison with other research methods

Method	Model	Input	Accuracy
Singh	LSTM	Raw signal	88.1%
Sellami	1D CNN	Raw signal	88.34%
Proposed study	LSTM	Raw signal	90.57%
	1D CNN-LSTM	Raw signal	92.3%

V. 결론

본 연구에서는 심전도 신호의 지역적인 특징을 추출하고 순환신경망의 장기 의존성 문제를 해결하기 위해 CNN-LSTM 하이브리드 결합 모델을 이용한 부정맥 분류 방법을 제안하였다.

이를 위해 먼저 잡음을 제거한 ECG 신호에서 R파를 검출하고 단일 비트 세그먼트를 추출하였다. 이후 부정맥 신호의 특징을 세밀하게 추출하도록 8개의 합성곱 계층으로 구성하고 이를 LSTM의 입력으로 사용한 후 가중치를 학습시키고 검증 데이터로 모델을 평가한 후 정상 및 부정맥 분류의 변화를 확인하였다. 제안한 방법의 타당성 검증을 위해 MIT-BIH 부정맥 데이터베이스를 사용하여 정확도(accuracy), 정밀도(precision), 재현율(recall), F1 score가 사용하였다. 성능평가 결과, 이상 패턴 검출률은 90%의 이상의 성능으로 제안하는 모델이 다른 모델 대비 우수한 성능을 보였다. 하지만 시간적 효율성 측면에서는 제안 모델의 연산이 복잡하여 향후 연구에서는 성능에 대한 최적화가 고려되어야 할 것으로 판단된다.

대상자의 증상 유무 및 종류에 따라 정확한 진단을 해야 하는 IoT 스마트 헬스케어 시스템의 경우 본 연구에서 제안한 하이브리드 결합모델은 부정맥 자동 분류 알고리즘에 효율적일 것으로 판단된다. 부정맥 분류 알고리즘에 대한 많은 연구가 진행되어 왔지만, 심장 질환자들에 대한 전문의의 해석을 정리 보관한 데이터베이스의 부재를 들 수 있다. 따라서 국내에 적합한 부정맥 분류 알고리즘을 정착시키기 위해서는 심장전문의와의 협력과 임상 데이터를 통한 데이터베이스 개발과 이를 통한 진단 알고리즘의 정착이 필요하다.

향후 연구에서는 부정맥 분류 자동화를 위한 연산 복잡도와 성능 최적화에 대한 연구가 고려되어야 한다.

ACKNOWLEDGEMENT

This work was supported by a 2-Year Research Grant of Pusan National University

REFERENCES

[1] H. M. Tun, W. K. Moe, and Z. M. Naing, "Analysis on conversion process from paper record ECG to computer based ECG," *MedCrave Online Journal of Applied Bionics and Biomechanics*, vol. 1, no. 2, pp. 69-81, Sep. 2017.

[2] A. Hossain, R. Quaresma, and H. Rahman, "Investigating factors influencing the physicians' adoption of electronic health record (EHR) in healthcare system of Bangladesh: An empirical study," *International Journal of Information Management*, vol. 44, pp. 76-87, Feb. 2019.

[3] Q. Qin, J. Li, Y. Yue, and C. Liu, "An Adaptive and Time-Efficient ECG R-Peak Detection Algorithm," *Journal of Healthcare Engineering*, vol. 2017, pp. 1-14, Sep. 2017.

[4] M. J. Goldman, *Principles of Clinical Electrocardiography*, Los Altos California: Lange Medical Pubns, 1986.

[5] K. Hanbay, "Deep neural network based approach for ECG classification using hybrid differential features and active learning," *Institution of Engineering and Technology*, vol. 13, no. 2, pp. 165-175, May. 2019.

[6] W. Li, "Deep Intermediate Representation and In-Set Voting Scheme for Multiple-Beat Electrocardiogram Classification," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6895-6904, Apr. 2019.

[7] M. Altuve and F. Hernández, "Multiclass Classification of Cardiac Rhythms on Short Single Lead ECG Recordings using Bidirectional Long Short-Term Memory Networks," *IEEE Latin America Transactions*, vol. 19, no. 7, pp. 1207-1216, Jul. 2021.

[8] I. S. Cho and H. S. Kwon, "Optimal Threshold Setting Method for R Wave Detection According to The Sampling Frequency of ECG Signals," *Journal of Korea Institute of Information and Communication Engineering*, vol. 21, no. 7, pp. 1420-1428, Jul. 2017.

[9] Y. Wei, J. Zhou, Y. Wang, Y. Liu, Q. Liu, J. Luo, C. Wang, F Ren, and L. Huang, "A Review of Algorithm & Hardware Design for AI-Based Biomedical Applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 2, pp. 145-163, Apr. 2020.

[10] S. Xu, J. Li, K. Liu, and L. Wu, "A Parallel GRU Recurrent Network Model and its Application to Multi-Channel Time-Varying Signal Classification," *IEEE Access*, vol. 7, pp. 118739-118748, Aug. 2019.

[11] Y. Li, Z. Xia, and Y. Zhang, "Standalone Systolic Profile Detection of Non-Contact SCG Signal With LSTM Network," *IEEE Sensors Journal*, vol. 20, no. 6, pp. 3123-3131, Mar. 2020.

[12] R. He, Y. Liu, K. Wang, N. Zhao, Y. Yuan, Q. Li, and H. Zhang, "Automatic Cardiac Arrhythmia Classification Using Combination of Deep Residual Network and Bidirectional LSTM," *IEEE Access*, vol. 7, pp. 102119-102135, Jul. 2019.

[13] S. Singh, S. K. Pandey, U. Pawar, and R. R. Janghel, "Classification of ECG Arrhythmia using Recurrent Neural Networks," *Procedia Computer Science*, vol. 132, pp. 1290-1297, Jun. 2018.

[14] A. Sellami and H. Hwang, "A robust deep convolutional neural network with batch-weighted loss for heartbeat classification," *Expert Systems with Applications*, vol. 122, pp. 75-84, May. 2019.



조익성(Ik-Sung Cho)

2003.02 : 한국해양대학교 공학박사
 2014.02 : 부산대학교 바이오정보전자공학과 공학박사
 2014.03 ~ 2018. 03 : 경운대학교 항공정보통신공학과 조교수
 2018.03 ~ 현재 : 대구대학교 창조융합학부 조교수
 ※관심분야 : ECG Preprocessing, 부정맥 분류 알고리즘



권혁송(Hyeog-Soong Kwon)

1995.08 : 영남대학교 전자공학과 공학박사
 1996.04-2006.02 : 밀양대학교 정보통신공학과 부교수
 2002.02~2003.02 : 미국 OSU 방문교수
 2011.07~2012.07 : 미국 조지아텍 방문교수
 2006.03 - 현재 : 부산대학교 IT 응용공학과 정교수
 ※관심분야 : 생체신호처리, ECG 특징점 추출, CDMA, 통신시스템