

Proposal of a Convolutional Neural Network Model for the Classification of Cardiomegaly in Chest X-ray Images

Min-Jeong Kim¹, Jung-Hun Kim^{2*}

¹Department of Biomedical Engineering, Kyungpook National University

²Bio-Medical Research institute, Kyungpook National University Hospital

Received: August 19 2021. Revised: October 20, 2021. Accepted: October 31, 2021.

ABSTRACT

The purpose of this study is to propose a convolutional neural network model that can classify normal and abnormal(cardiomegaly) in chest X-ray images. The training data and test data used in this paper were used by acquiring chest X-ray images of patients diagnosed with normal and abnormal(cardiomegaly). Using the proposed deep learning model, we classified normal and abnormal(cardiomegaly) images and verified the classification performance. When using the proposed model, the classification accuracy of normal and abnormal(cardiomegaly) was 99.88%. Validation of classification performance using normal images as test data showed 95%, 100%, 90%, and 96% in accuracy, precision, recall, and F1 score. Validation of classification performance using abnormal(cardiomegaly) images as test data showed 95%, 92%, 100%, and 96% in accuracy, precision, recall, and F1 score. Our classification results show that the proposed convolutional neural network model shows very good performance in feature extraction and classification of chest X-ray images. The convolutional neural network model proposed in this paper is expected to show useful results for disease classification of chest X-ray images, and further study of CNN models are needed focusing on the features of medical images.

Keywords: Convolutional Neural Network, Deep learning, Chest X-ray, Classification

I. INTRODUCTION

심장비대(Cardiomegaly)는 심장 내부 공간의 확장 또는 심실 비대에 의해 심장이 흉곽의 내경보다 50% 이상 커진 상태를 의미하며, 심혈관계의 합병증을 유발하여 임상적으로 문제가 되는 질환이다^[1-3]. 심장비대는 심전도검사(Electrocardiogram, ECG), 단순 흉부 방사선 검사(X-ray), 심장 초음파 검사(Echocardiography) 등을 이용하여 조기 발견하여 다른 합병증이 유발되는 것을 막는 것이 매우 중요하다. 하지만 의료 영상을 판독하는 의사에 따라 숙련도가 다르고 같은 의사가 판독하더라도 검사 당시의 피로도 등에 따라 결과에 차이가 발생할 위험성이 존재한다^[4-7].

최근 많은 연구가 진행되고 있는 심층 인공 신경망(Deep Neural Network)은 입력 데이터로부터 다양한 특징을 추출하고 계층적으로 학습하여 뛰어난 성능을 이끌어내어 다양한 응용 분야에서 적용되고 있다^[8-11]. 특히, 영상 인식과 분류 등의 응용 분야에서 뛰어난 성능을 보여주고 있는 합성곱 신경망 모델(Convolutional Neural Network)은 의료 영상 분야에서도 전문가 수준의 성능을 보여주어 질환 진단 및 분류에도 활용되고 있다^[12-15].

흉부 X-선 영상 분석 분야의 딥러닝 (Deep Learning) 활용 예시로는, 흉부 X-선 영상을 사용하여 무기폐(Atelectasis), 심장비대(Cardiomegaly), 흉수(Effusion), 폐 침윤(Infiltration), 폐 종괴(Mass), 폐 결절(Nodule), 폐렴(Pneumonia), 폐 기흉(Pneumothorax) 등의 병변을 판단하는 알고리즘을 제안한 Yao^[16]

* Corresponding Author: Jung-Hun Kim

E-mail: fainal2@naver.com

Tel: +82-053-950-7956

등, Wang^[17] 등이 대표적이다. 또한, Islam^[18] 등은 AlexNet^[9], VGGNet^[10], ResNet^[11] 등 다양한 심층 인공 신경망(Deep Neural Network)에서 흉부 X-선 영상을 이용하였을 때의 분류 성능을 비교하는 연구를 진행하였고, 모든 신경망에서 85% 이상의 분류 정확도를 보였다. Que^[19] 등은 U-Net^[20]을 기반으로 한 합성곱 신경망 모델을 이용하여 흉부 X-선 영상에서 심장비대의 유무를 분류하는 연구를 진행하였고, 93.75%의 분류 정확도를 얻었다.

하지만 자연 영상 분류 또는 인식에 사용되는 학습 데이터와 달리 의료 영상은 수집하기에 어려움이 있어 학습 데이터의 수가 부족하다. 이러한 이유로 대부분의 의료 영상 연구가 일반적인 자연 영상 데이터를 기반으로 학습한 합성곱 신경망 모델을 이용하여 학습이 진행되고 있다. 하지만 의료 영상은 일반적인 자연 영상과 달리 RGB 값이 아닌 그레이스케일 흑백으로 구성되어 있고 인체 내부를 표현하고 있다는 점에서 그레이스케일 흑백의 의료 영상과 같은 특수한 분야의 영상 특징을 잘 반영할 수 있는 모델의 개발이 필요하다.

본 논문에서는 흉부 X-선 영상에서 정상 심장과 비정상 심장(심장비대)을 분류할 수 있는 합성곱 신경망 모델을 제안하고자 한다. 제안하는 알고리즘은 16개의 레이어를 가지는 합성곱 신경망으로 정상 심장과 심장비대가 의심되는 비정상 심장으로 분류한다. 또한, 제안된 합성곱 신경망 모델을 이용하여 정상 심장과 비정상 심장(심장비대)을 분류하고, 분류된 영상이 실제로 정답과 얼마나 가까운지 검증하고자 하였다.

II. MATERIAL AND METHOD

1. 개발 및 구현 환경

본 연구에서는 딥러닝 모델링을 구현하기 위해 Python(version 3.8.8) 언어 기반의 통합 개발 환경인 주피터 노트북(Jupyter Notebook)을 사용하였다. 딥러닝 학습을 위한 시스템은 Intel (R) Core (TM) i7-8700k(Intel, Santa Clara, CA, USA), NVIDIA GeForce GTX 1080 Ti(NVIDIA, Santa Clara, CA, USA), Tensorflow 2.3.0, Keras 2.4.3을 사용하였다.

2. 데이터 수집

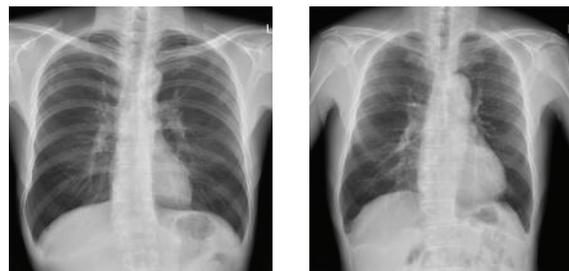
연구에 사용된 데이터는 경북대학교병원 IRB(Institutional Review Board, 의학 연구 윤리 심의위원회)의 심의 절차에 따라 승인을 받아 수집을 진행하였다(IRB 번호: KNUH 2020-07-018-002). 2010년 1월부터 2020년 12월까지 경북대학교병원에 내원한 만 20세 이상에서 95세 미만의 환자들의 흉부 X-선 이미지를 획득하였다. 수집한 이미지 데이터는 정상 심장 526장, 비정상 심장(심장비대) 500장으로 구성되었다. 분류를 위한 데이터의 구성은 Table 1과 같이 전체 데이터의 80%를 학습 데이터에 사용하였으며, 나머지 20%를 최종 모델을 검증하는 테스트 데이터 세트로 구성하였다.

Table 1. Classification of Data

Components	Normal	Abnormal
Training data	226	250
Validation data	200	150
Test data	100	100

3. 데이터 Annotation

본 연구에서 사용된 데이터는 흉부 방사선 전문의가 각각 진단하여 Fig. 1과 같이 정상(Normal) 이미지와 비정상 심장(Abnormal)으로 구분하였다. 각 데이터의 Annotation은 AVIEW(Coreline Europe GmbH, Eschborn, Germany)을 활용하여 전문의가 흉부 X-선 영상에서 심장 크기에 맞추어 직사각형 모양의 외곽선을 그리는 방식으로 진행되었다.



(A) Normal

(B) Abnormal

Fig. 1. Chest X-ray Images.

4. 제안하는 합성곱 신경망 모델 구조

최근 합성곱 신경망에 관한 연구가 많이 진행되면서 합성곱 신경망 모델의 분류 성능을 향상시키기 위한 다양한 방법이 제안되고 있다. 이에 본 연구에서는 Fig. 2와 같이 총 16개의 레이어로 구성된 합성곱 신경망을 제안하고자 한다.

본 연구에서 제안하는 합성곱 신경망 모델은 8개의 Convolution 레이어와 6개의 Fully connected 레이어, 입력 영상을 서브 샘플링(Subsampling) 하여 크기를 축소시키는 2개의 Max pooling 레이어로 구성되어 있다. 모든 Convolution 레이어의 필터 크기는 파라미터의 수와 학습 소요 시간을 줄이기 위해 2 x 2로 고정하였다.

첫 번째에서 네 번째 레이어는 2 x 2 크기를 가지는 160개의 합성곱 커널로 구성되어 있다. 다섯 번째 레이어는 2 x 2 크기를 가지는 Max pooling 레이어로 입력 영상의 크기를 절반으로 줄여 14 x 14 크기의 영상을 출력한다. 여섯 번째에서 아홉 번째 레이어는 2 x 2 크기를 가지는 320개의 합성곱 커널로 구성되어 있다. 이어서 열 번째는 2 x 2 크기를 가지는 Max pooling 레이어로 구성되어 있다. 이러한 Convolution 레이어와 Max pooling 레이어는 입력된 흉부 X-선 이미지에서 정상과 비정상적으로 분류하기 위해 사용되는 특징들을 추출하고, 학습 과정을 통해서 더 좋은 특징을 선정하게 된다.

열한 번째에서 열여섯 번째 레이어는 이전의 레이어의 모든 뉴런들과 연결되어있는 Fully connected 레이어로 구성되어 있다. 열한 번째 레이어는 선정된 특징을 Fully connected 레이어에 전달하기 위해 2차원 데이터를 1차원 데이터로 바꾸어 주는 Flatten 레이어로 구성되어 있다. 열두 번째에서 열여섯 번째 레이어는 Fully connected 레이어로 각각 8001개, 512개, 256개, 128개의 노드로 구성되어 있으며, 앞의 학습 과정을 통하여 선정된 특징을 바탕으로 정상 심장과 비정상 심장(심장비대)을 분류한다. 그리고 마지막 출력 레이어는 입력 영상이 정상 심장 혹은 비정상 심장(심장비대)인지에 관한 결과를 출력하는 1개의 노드로 구성된 Sigmoid 레이어로 구성되어 있다.

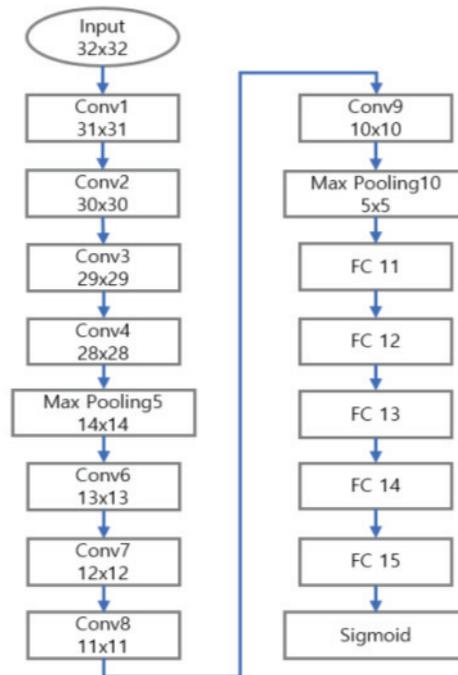


Fig. 2. A Flowchart of the Layer Constituting the Cnn Model for Normal and Abnormal(Cardiomegaly) Classification.

5. 데이터 Augmentation 방법

데이터 증대(Augmentation)는 영상 인식 분야에서 제한된 데이터를 이용하여 충분한 학습을 진행해야 할 때 사용된다. 데이터 증대는 입력 영상에 Translation과 같은 연산을 수행하여 진행한다.

본 연구에서는 데이터 증대 기법을 사용하여 이미지를 시계 및 반시계 방향으로 0°~90° 이내의 범위로 회전시키고, 너비와 높이 및 이미지 전체를 60%~140%로 확대 및 축소하는 과정을 랜덤하게 적용하여 이미지 수를 증가시켰다.

위와 같은 방법을 이용하여 테스트 데이터 세트는 제외하고 학습 데이터 세트에 대해서만 증대를 진행하였다. 정상 심장과 비정상 심장(심장비대)에서 30배 증대시킨 학습 데이터 세트의 개수는 각각 14,227장, 26,251장이며 60배 증대시킨 학습 데이터 세트의 개수는 각각 28,727장, 31,771장이다. 증대시킨 데이터를 사용하여 학습 데이터의 개수에 따른 학습 및 테스트 정확도를 비교하고자 하였다.

Table 2. 30 Times and 60 Times Augmented Data

	30 times augmentation		60 times augmentation	
	Normal	Abnormal	Normal	Abnormal
Training data (Validation data)	14,227 (200)	16,251 (150)	28,727 (200)	31,771 (150)
Test data	200	150	200	150
Total	14,627	16,551	31,127	32,001

6. 모델 최적화 방법

본 논문에서는 학습을 위한 최적화 알고리즘 (Optimizer)으로 Adam(Adaptive Moment Estimation)을 사용하였고, 손실 함수로는 BinaryCrossentropy를 사용하였다^[21]. 배치 사이즈(Batch size)는 데이터 수가 많아 학습을 진행하는 데 많은 시간이 소요되어 효과적인 학습 속도를 위해 64로 지정하였다. 에폭 (Epoch)은 초기에 50으로 설정하였으나 일정한 시점 이후로 정확도가 더 증가하지 않고 유지되거나 감소하는 현상이 발생하여 20으로 지정하였다. 학습률(Learning rate)은 6가지로 설정하여 학습률에 따른 분류 정확도를 비교하고자 하였다. 학습에 사용된 6가지의 구체적인 학습률은 0.005, 0.003, 0.001, 0.0005, 0.0003, 0.0001이다.

7. 모델 평가

본 연구에서는 학습을 진행한 모델들의 성능을 테스트 데이터 세트를 사용하여 비교하였다. 모델 성능 비교에는 일반적으로 널리 사용되는 정량적 지표인 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1 Score을 사용하여 나타냈다^[22].

정확도(Accuracy)는 전체 흉부 X-선 영상에서 정상 심장 또는 비정상 심장(심장비대)으로 분류된 클래스에서 얼마나 많은 클래스가 올바른 라벨을 가졌는지에 대한 확률을 나타낸다. 정밀도(Precision)는 분류된 정상 심장 데이터 집합에서 실제 정상 심장 데이터가 얼마나 존재하는지 예측 능력을 평가한 지표이다. 재현율(Recall)은 sensitivity로도 불리며, 실제 정상 심장 데이터에서 얼마나 많은 데이터가 정상 심장 데이터로 분류되는지에 대한 확률을 나타내는 지표이다. F1 Score는 정밀도와 재현율의 조화평균으로 데이터 라벨이 불균형을 이루는 경우, 모델의 성능을 하나의 숫자로

표현하여 정확하게 평가할 수 있다. 성능 지표에 대한 계산식을 아래의 Eq. (1) ~ (4)에 작성하였다.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \times 100 \quad (3)$$

$$F1 \text{ Score} = \frac{2 * Precision * Recall}{Precision + Recall} \times 100 \quad (4)$$

- TP : 진양성(true positive)
- FP : 위양성(false positive)
- TN : 진음성(true negative)
- FN : 위음성(false negative)

III. RESULT

1. 데이터 증대에 따른 정확도 결과

본 논문에서는 데이터의 증대 여부와 증대된 배수에 따라 정확도에 미치는 영향을 확인하고자 하였다. 데이터 증대를 진행하지 않은 데이터를 이용하여 정상 심장과 비정상 심장(심장비대)을 분류하였을 때의 학습 및 테스트 정확도는 92.44%, 92.0%를 보여주었다. 30배 데이터 증대를 진행한 데이터를 이용하였을 때의 학습 및 테스트 정확도는 99.71%, 88.0%를 보여주었고, 60배 데이터 증대를 진행한 데이터를 이용하였을 때는 99.84%, 95.0%의 정확도를 보여주었다. Table 3에 데이터 증대에 따른 정확도 결과를 정리하여 나타냈다. 60배 데이터 증대를 진행하였을 때, 가장 높은 학습 및 테스트 정확도를 보여주었다.

Table 3. Accuracy(%) according to Augmented Data

Data	30 times augmentation	60 times augmentation
Train Accuracy	92.44	99.84
Test Accuracy	92.0	95.0

2. 학습률 변화에 따른 정확도 결과

본 논문에서는 학습률(Learning rate)은 6가지로

설정하여 학습률에 따른 정확도를 비교하고자 하였다. 이때, 학습에 사용된 데이터는 데이터 증대에 따른 정확도 비교에서 가장 높은 정확도를 보여주었던 60배로 증대한 데이터 세트를 사용하였다. 학습률에 따른 학습 정확도를 0.005, 0.003, 0.001, 0.0005, 0.0003, 0.0001 순서로 살펴본다. 학습 정확도는 차례로 52.52%, 52.52%, 52.52%, 52.52%, 99.71%, 99.88%이다. 학습률이 0.0001일 때, 99.88%로 가장 높은 정확도를 보여주었다. Fig. 3은 학습률에 따른 학습 성능을 보여주는 그래프이다.

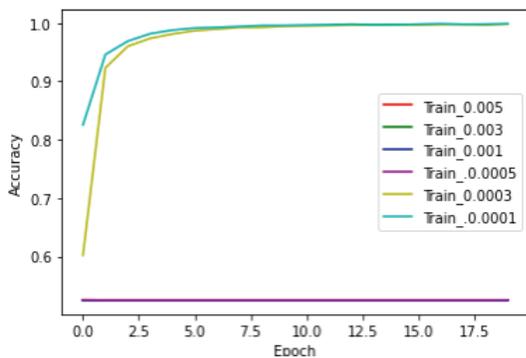


Fig. 3. Compare Model Accuracy between Learning Rates.

3. 제안한 모델의 정확도 결과

본 논문에서는 제안하는 합성곱 신경망 모델을 이용하여 흉부 X-선 영상의 정상 심장과 비정상 심장(심장비대) 분류에 대한 정확도를 확인하고 분류 성능을 평가하고자 하였다. 학습을 진행하기 위해 데이터 증대에 따른 정확도 비교에서 가장 높은 정확도를 보여주었던 60배로 증대한 데이터 세트를 사용하였고, 학습률에 따른 정확도 비교에서 가장 높은 정확도를 보여주었던 0.0001로 설정하여 학습을 진행하였다.

본 논문에서 제안한 합성곱 신경망 모델의 정상 심장과 비정상 심장(심장비대) 분류에 대한 학습 정확도의 경우 99.84%, 검증 정확도의 경우 94%의 결과를 보여주었다. 학습 손실 값의 경우 0.0051, 검증 손실 값의 경우 0.2628의 결과를 보여주었다. Fig. 4는 제안한 모델에서의 정상 심장과 비정상 심장(심장비대)의 분류 결과를 보여주는 그래프이다.

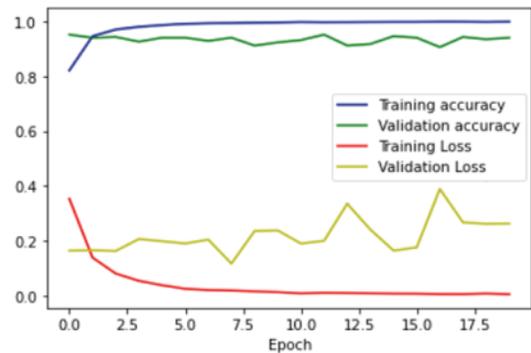


Fig. 4. Classification Result of the Proposed Model.

본 논문에서 제안한 합성곱 신경망 모델의 분류 성능을 검증하기 위하여 테스트 데이터 세트를 사용하여 정확도, 정밀도, 재현율, F1 Score를 구하고 평가하였다. 테스트 데이터에 대한 성능 평가 지표들은 Eq. (1) ~ (4)를 통하여 계산하였다.

정상 심장 데이터와 비정상 심장(심장비대) 데이터를 사용하였을 때의 제안한 합성곱 신경망 모델의 분류 성능을 정확도, 정밀도, 재현율, F1 Score 순서로 살펴본다. 정상 심장의 경우 95%, 100%, 90%, 96%이고, 심장비대의 경우 95%, 92%, 100%, 96%이다. Table 4에서 테스트 데이터 세트에 대한 성능 결과를 확인할 수 있다.

Table 4. Result of Classification using Test Data

	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Normal	95.0	100.0	90.0	96.0
Abnormal	95.0	92.0	100.0	96.0

IV. DISCUSSION

본 논문에서는 의료 영상 분야에서 최근 많이 활용되고 있는 합성곱 신경망(Convolution Neural Network)을 기반으로 딥러닝 모델을 제안하여 흉부 X-선 영상 진단에 도움을 주고자 하였다. 이를 위해 정상 심장과 비정상 심장(심장비대)을 분류하는 모델을 본 연구에서 제시하였다.

그 결과 학습 정확도의 경우 99.84%, 검증 정확도의 경우 94%로 높은 수준의 정확도를 보여주었다. 또한, 학습 손실 값의 경우 0.0051, 검증 손실 값의 경우 0.2628로 0에 수렴하는 낮은 손실 값을

보여주었다. 이러한 분류 결과는 제안하는 합성곱 신경망 모델이 정상 심장 흉부 X-선 영상과 비정상 심장(심장비대) 흉부 X-선 영상에 대한 분류에서 매우 우수한 성능을 보여주고 있다고 판단된다.

본 연구에서 데이터 증대 여부와 학습률의 변화에 따라 제안된 모델을 활용하여 결과를 획득하였다. 데이터 증대 여부에 따라서는 학습 데이터를 증대하는 배수가 높아질수록 학습 및 테스트 정확도가 증가하는 추이를 보여주었다. 학습률의 변화에 따라서는 학습률이 감소할수록 정확도가 향상되고 안정적으로 학습이 진행되었다. 이는 많은 학습 데이터를 사용할수록 충분한 학습이 진행되어 높은 결과를 얻을 수 있었던 것으로 생각된다.

테스트 데이터를 이용한 성능 평가에서는 정확도, 정밀도, 재현율, F1 Score 순서로 살펴보았을 때, 정상 심장은 95%, 100%, 90%, 96%이고, 비정상 심장(심장비대)은 95%, 92%, 100%, 96%로 모두 높은 분류 성능을 보여주었다.

앞서 싱가포르의 연구팀이 U-Net 기반의 합성곱 신경망 모델을 이용하여 정상 심장과 비정상 심장(심장비대)을 정확도 93.75%로 분류한 사례가 있다. 성능 평가에서는 정밀도 100%, 재현율 89.29%, F1 Score 94.34%를 보여주었다. 데이터 측면에서 보았을 때, 본 연구가 싱가포르 연구팀보다 데이터 증대를 통해 많은 데이터를 사용하여 상대적으로 충분한 학습이 진행되었다고 판단된다. 또한, 방법론적 측면에서는 본 연구에서 제안하는 합성곱 신경망 모델이 의료 영상인 흉부 X-선 영상 특징을 잘 반영할 수 있도록 개발되었기 때문에 위 연구보다 더 높은 분류 성능 결과를 얻을 수 있었다고 생각된다. 이러한 결과는 제안하는 합성곱 신경망 모델이 흉부 X-선 영상에서 정상 심장과 심장비대가 의심되는 비정상 심장 분류에서 매우 우수한 성능을 보여주고 있다고 판단된다.

V. CONCLUSION

본 논문에서는 흉부 X-선 영상의 정상 심장과 비정상 심장(심장비대)을 분류할 수 있는 합성곱 신경망 모델을 제안하였다. 제안한 모델은 우수한 성능을 보여주었으며 흉부 X-선 영상의 질환 분류에

있어 유용한 결과를 보여줄 것으로 판단된다. 흉부 X-선 영상에 딥러닝을 적용하려는 여러 연구에 파라미터에 따른 정보 제공에 도움이 될 것으로 생각된다. 향후 본 논문의 결과가 다른 의료 영상에서도 일반적으로 확인되는 특징인지 알아보기 위하여 더 다양하고 많은 데이터를 가진 추가 연구가 이루어져야 할 것이며, 결과를 바탕으로 흉부 X-선 영상과 같이 특수한 영상에 알맞은 모델의 개발이 지속적으로 이루어져야 할 것이다.

Reference

- [1] C. S. Danzer, "The cardi thoracic ratio: an index of cardiac enlargement", *The American Journal of the Medical Sciences*, Vol. 157, No. 4, pp. 1827-1924, 1919.
- [2] S. H. Paek, J. M. Lee, S. J. Han, Y. W. Bahk, "Evaluation of cardiac measurements in healthy Korean adults", *Journal of the Korean radiological society*, Vol. 14, No. 1, pp. 57-62, 1978.
- [3] Y. S. Kim, H. J. Park, S. H. Park, H. J. Chun, B. G. Choi, "A CT criteria of cardiomegaly", *Journal of the Korean Radiological Society*, Vol. 57, No. 3, pp. 235-238, 2007.
- [4] S. H. Yang, J. S. Lee, C. S. Kim, "The Accuracy of Echocardiography and ECG in the Left Ventricular Hypertrophy", *The Journal of the Korea Contents Association*, Vol. 16, No. 2, pp. 666-672, 2016. <http://dx.doi.org/10.5392/JKCA.2016.16.02.666>
- [5] E. D. Frohlich, "Left ventricular hypertrophy as a risk factor", *Cardiology clinics*, Vol. 4, No. 1, pp. 137-144, 1986.
- [6] H. Amin, W. J. Siddiqui, "Cardiomegaly", *StatPearls [internet]*, 2020.
- [7] S. S. Alghamdi, I. Abdelaziz, M. Albadri, S. Alyanbaawi, R. Aljondi, A. Tajaldeen, "Study of cardiomegaly using chest x-ray", *Journal of Radiation Research and Applied Sciences*, Vol. 13, No. 1, pp. 460-467, 2020. <http://dx.doi.org/10.1080/16878507.2020.1756187>
- [8] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning, *Nature*", Vol. 521, No. 7553, 436-444, 2015.
- [9] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet

- classification with deep convolutional neural networks", *Advances in neural information processing systems*, 25, 1097-1105, 2012.
- [10] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition", arXiv preprint arXiv:1409.1556, 2014.
- [11] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016.
- [12] M. H. Yap, M. Goyal, F. Osman, E. Ahmad, R. Martí, E. Denton, A. Juette, R. Zwiggelaar, "End-to-end breast ultrasound lesions recognition with a deep learning approach", In *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*, International Society for Optics and Photonics, Vol. 10578, pp. 1057819, 2018. <https://doi.org/10.1117/12.2293498>
- [13] S. Han, H. K. Kang, J. Y. Jeong, M. H. Park, W. Kim, W. C. Bang, Y. K. Seong, "A deep learning framework for supporting the classification of breast lesions in ultrasound images", *Physics in Medicine & Biology*, Vol. 62, No. 19, pp. 7714, 2017. <http://dx.doi.org/10.1088/1361-6560/aa82ec>
- [14] B. Sahiner, A. Pezeshk, L. M. Hadjiiski, X. Wang, K. Drukker, K. H. Cha, R. M. Summers, M. L. Giger, "Deep learning in medical imaging and radiation therapy", *Medical physics*, Vol. 46, No. 1, pp. e1-e36, 2019. <http://dx.doi.org/10.1002/mp.13264>
- [15] S. Lim, M. Lee, "A Computer Aided Diagnosis Algorithm for Classification of Malignant Melanoma Based on Deep Learning", *Journal of the Korea Society of Digital Industry and Information Management*, Vol. 14, No. 4, pp. 69-77, 2018.
- [16] L. Yao, E. Poblens, D. Dagunts, B. Covington, D. Bernard, K. Lyman, "Learning to diagnose from scratch by exploiting dependencies among labels", arXiv preprint arXiv:1710.10501, 2017.
- [17] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases", In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2097-2106, 2017.
- [18] M. T. Islam, M. A. Aowal, A. T. Minhaz, K. Ashraf, "Abnormality detection and localization in chest x-rays using deep convolutional neural networks", arXiv preprint arXiv:1705.09850, 2017.
- [19] Q. Que, Z. Tang, R. Wang, Z. Zeng, J. Wang, M. Chua, T. S. Gee, X. Yang, B. Veeravalli, "CardioXNet: automated detection for cardiomegaly based on deep learning", In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 612-615, 2018.
- [20] O. Ronneberger, P. Fischer, T. Brox, "U-net: Convolutional networks for biomedical image segmentation", In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234-241, 2015, Springer, Cham.
- [21] D. P. Kingma, J. Ba, "Adam: A method for stochastic optimization", arXiv preprint arXiv:1412.6980, 2014.
- [22] N. Japkowicz, M. Shah, "Evaluating learning algorithms: a classification perspective", Cambridge University Press, 2011.

흉부 X-선 영상에서 심장비대증 분류를 위한 합성곱 신경망 모델 제안

김민정¹, 김정훈^{2,*}

¹경북대학교대학원 의용생체공학과

²경북대학교병원 생명 의학 연구원

요 약

본 논문에서는 흉부 X선 영상에서 정상 심장과 비정상 심장(심장비대)을 분류할 수 있는 합성곱 신경망 모델을 제안하고자 한다. 학습 및 테스트 데이터로는 경북대학교병원에 내원하여 정상과 심장비대를 진단 받은 환자들의 흉부 X-선 이미지를 획득하여 사용하였다. 제안된 합성곱 신경망 모델을 이용하였을 때의 정상 심장 및 비정상 심장(심장비대) 분류 정확도는 99.88%였다. 정상 심장 영상을 테스트 데이터로 사용하였을 때의 정확도, 정밀도, 재현율 및 F1 Score는 95%, 100%, 90%, 96%였다. 비정상 심장(심장비대) 영상을 테스트 데이터로 사용하였을 때의 정확도, 정밀도, 재현율 및 F1 Score는 95%, 92%, 100% 및 96%였다. 이러한 학습 및 테스트 분류 결과로 제안된 합성곱 신경망 모델은 흉부 X-선 영상의 특징 추출 및 분류에서 매우 우수한 성능을 보여주고 있다고 판단된다. 본 논문에서 제안하는 합성곱 신경망 모델은 흉부 X-선 영상의 질한 분류에 있어 유용한 결과를 보여줄 것으로 판단되며, 다른 의료 영상에서도 동일한 결과를 나타내는지 알아보기 위하여 추가적인 연구가 이루어져야 할 것이다.

중심단어: 합성곱 신경망, 딥러닝, 흉부 X-선, 분류

연구자 정보 이력

	성명	소속	직위
(제1저자)	김민정	경북대학교대학원 의용생체공학과	석사과정
(교신저자)	김정훈	경북대학교병원 생명 의학 연구원	교수