

<중설>

AI 의료영상 분석의 개요 및 연구 현황에 대한 고찰

홍준용¹⁾·박상현¹⁾·정영진²⁾¹⁾동서대학교 융합방사선학과·²⁾동서대학교 방사선학과

Artificial Intelligence Based Medical Imaging: An Overview

Hong Jun-Yong¹⁾·Park Sang Hyun¹⁾·Jung Young-Jin²⁾¹⁾Department of Multidisciplinary Radiological Science, Dongseo University²⁾Department of Radiology, Dongseo University

Abstract Artificial intelligence(AI) is a field of computer science that is defined as allowing computers to imitate human intellectual behavior, even though AI's performance is to imitate humans. It is grafted across software-based fields with the advantages of high accuracy and speed of processing that surpasses humans. Indeed, the AI based technology has become a key technology in the medical field that will lead the development of medical image analysis. Therefore, this article introduces and discusses the concept of deep learning-based medical imaging analysis using the principle of algorithms for convolutional neural network(CNN) and back propagation. The research cases application of the AI based medical imaging analysis is used to classify the various disease(such as chest disease, coronary artery disease, and cerebrovascular disease), and the performance estimation comparing between AI based medical imaging classifier and human experts.

Key Words: Artificial Intelligence, Medical Imaging, Deep Learning, Artificial Neural Network

중심 단어: 인공지능, 의료영상, 딥 러닝, 인공지능경망

I. 서 론

인공지능(artificial intelligence; AI)이란 컴퓨터 공학의 한 분야로서 컴퓨터로 하여금 인간의 지적 행동을 모방하도록 하는 것으로 정의할 수 있는데, 비록 인공지능의 작업수행이 인간을 모방하는 것일지라도 방대한 양의 데이터 학습을 통해 구축된 높은 정확성과 인간을 뛰어넘는 빠른 처리 속도를 장점으로 앞세워 소프트웨어 기반 분야의 전반에 걸쳐 접목되고 있다[1]. 이는 의료 분야에서도 예외는 아니며, 특히 AI 의료영상 분석 기술은 의료영상 분야의 발전을 선도할 핵심 기술로 자리 잡고 있다[2].

AI 의료영상 분석이란 사전 학습된 AI가 의료영상으로부터

특징(feature)을 추출하여 병변을 진단하는 기술이다[3]. 의료영상에서 특징이란 명도(brightness), 대조도(contrast), 공간주파수(spatial-frequency), 균질성(homogeneity), 곡률(curvature), 길이(length) 등 영상의 데이터를 통해 정량적으로 나타낼 수 있는 것을 의미하는데, 각 병변은 서로를 구분 짓게 하는 고유의 특징을 가진다[4]. AI 의료영상 분석 기술은 특징을 추출하는 방법으로 최근 딥-러닝(deep learning)에 기반을 둔 학습모델이 주목을 받고 있다[5].

딥 러닝은 인간의 신경계를 모방한 인공지능경망(Artificial Neural Network; ANN)의 일종으로, 여러 층으로 쌓인 신경망 내부에서 각종 파라미터를 갱신하며 학습을 수행하는 기법이다. 딥 러닝을 이용하여 영상을 분석하는 개념은 일찍이

This work was supported by Dongseo University, "Dongseo Cluster Project" Research Fund of 2020 (DSU-2020004)

Corresponding author: Young-Jin Jung, Department of Radiological Science, Dongseo University, 47, Jurey-ro, Sasang-Gu, Busan, 47011, Republic of Korea / Tel: +82-51-320-2871 / E-mail: microbme@dongseo.ac.kr

Received 06 May 2020; Revised 07 June 2020; Accepted 22 June 2020

Copyright ©2020 by The Korean Journal of Radiological Science and Technology

1980년대에 소개되었으나[6], 당시의 느린 컴퓨터 연산 속도와 데이터 셋 확보의 어려움 등의 한계점이 있었다. 데이터 셋이란 딥 러닝이 특정 질환을 분류할 수 있도록 학습시키는데 사용되는 방대한 양의 의료영상 데이터 모음을 의미한다. 데이터 셋은 신뢰할 수 있는 전문가 집단에 의한 레이블링(질환명 등)을 포함해야 하며, 일부를 사전 학습에 사용하고 나머지를 성능 검증에 사용하는 것이 일반적이다. 과거에는 데이터 전송 속도 등의 문제로 데이터 셋의 공유가 적었고, 컴퓨터 연산 속도의 한계로 딥 러닝은 실용화하기 어려운 기술로 여겨졌다. 하지만, 최근에는 하드웨어의 발전으로 저비용 그래픽처리장치(Graphics Processing Unit; GPU)가 등장하여 복잡한 딥 러닝 연산에 소요되는 시간이 크게 단축되었고, 의료영상 저장전송시스템(picture archiving and communication system; PACS)과 인터넷의 발전으로 딥 러닝의 학습에 사용되는 방대한 규모의 데이터 셋 공유가 늘어났다. 또한, 과잉적합을 방지하는 다양한 규제기법이 개발됨에 따라 기존 딥 러닝 모델의 단점이 다수 극복되었다[7-10].

본 연구는 문헌적 고찰을 통해 인공지능의 이해와 영상의학적 접근에 대한 이해를 높이는 데 목적을 두고자 한다. 이에 의료영상 분석에서의 딥 러닝의 개요, 알고리즘의 원리, 그리고 연구 사례를 소개하고 논의한다. 2장에서는 인공지능, 머신러닝과 딥 러닝의 개념을 소개한다. 3장에서는 딥 러닝을 구현하기 위한 합성곱 신경망의 학습 원리와 최적화 과정을 소개한다. 4장에서는 다양한 질환의 분석을 위해 사용된 연구 사례를 소개하며, 그에 대한 고찰과 결론으로 글을 마무리한다.

II. AI 의료영상 분석의 개요

Russel, S의 저서 Artificial Intelligence: A Modern Approach에 따르면, 다음의 4가지 관점으로 인공지능을 소개하고 있다[11]. 첫째, 인간적으로 사고하는(Thinking Humanly), 둘째, 인간적으로 행동하는(Acting Humanly), 셋째, 이성적으로 사고하는(Thinking Rationally), 넷째, 이성적으로 행동하는(Acting Rationally) 의 다소 포괄적인 개념으로 정의된다. 이 개념을 실제 문제 해결에 적용하기 위해서는 알고리즘 기반의 시스템이 필요한데, 이를 인공지능의 하위 개념인 머신 러닝(Machine Learning; ML)이라고 한다(Fig. 1). 머신 러닝이란 어떠한 작업에 대해 꾸준한 학습을 통하여 작업의 성능을 높이는 것으로 정의된다[12]. 다시 말해, 의료영상의 분석이라는 작업에 대해 영상을 꾸준히 학습하여 분석의 정확도를 높이는 것이다. 의료영상 분야에서 머신 러닝의 학습과 분석에는 의료영상의 특징

(feature)을 사용한다. 특징이란 명도, 대조도, 공간주파수, 균질성, 곡률, 길이 등 영상의 데이터만을 통해 정량적으로 나타낼 수 있는 것을 의미한다. 따라서, 각 병변은 서로를 구분 짓게 하는 고유의 특징을 가질 것으로 여겨진다. 이러한 고유의 특징은 학습 과정에서 학습 모델에게 해당 병변의 정보를 제공한다[13]. 판독과정에서는 판독의 대상이 되는 영상의 특징을 추출하여 사전에 학습된 해당 병변의 정보와의 유사도를 산출하는 원리로 병변을 분류한다.

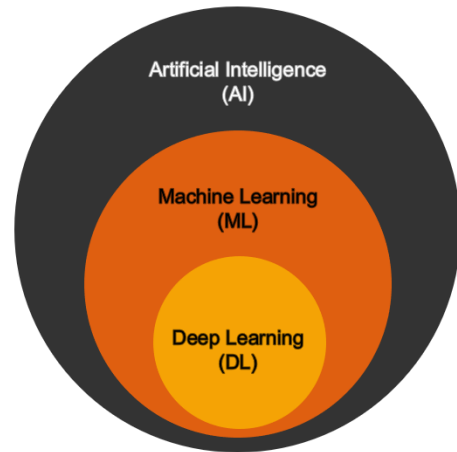


Fig. 1. Concept of artificial intelligence, machine learning, and deep learning

머신 러닝 기반 기법은 학습 방식에 따라 크게 지도 학습(supervised learning)과 비지도 학습(unsupervised learning)으로 나뉜다[14]. 지도학습은 정답을 알려주면서 학습을 시키는 과정으로 설명할 수 있는데, 예를 들어 종양 부위를 분할하는 AI 모델을 학습시키기 위해서 실제 의사가 직접 영상에 ROI를 표기(labeling)해야 한다. 알고리즘은 정답 값인 이 ROI를 기준으로 종양의 명도, 대조도, 균질성, 공간주파수, 곡률, 길이 등의 특징을 추출하여 학습한다. 따라서 이 과정에 많은 시간이 소요되고, ROI 선정에 있어 주관적 관점이 개입되므로 객관적인 데이터 분류로 받아들이기에는 어려움이 있다. 비지도 학습의 경우 정답을 알려주지 않고 데이터의 특징에 따라 군집화하는 과정으로 설명할 수 있는데, 종양 영역이 강도, 대조도, 균질성, 공간주파수, 곡률, 길이 등의 특징에 대해 정상 소견의 영상에서는 나타나지 않는 이상값(outlier)을 보이는 특성을 활용한 기법이다. 다시 말해, 종양 영역에 의한 이상값은 정상 소견의 영상에서 나타나지 않으므로, 이 차이를 이용해서 각 영상을 서로 다른 군집으로 분류할 수 있다. 하지만, 노이즈 신호에 의한 이상값에도 민감하게 반응한다는 단점이 있다.

머신 러닝의 하위 개념으로는 딥 러닝(Deep Learning;

DL)이 존재한다. 머신 러닝과 비교하여 딥 러닝 기반 기법의 가장 큰 차이점은 학습에 필요한 특징 추출 모델을 수동으로 제공할 필요가 없다는 것이다[15]. 딥 러닝은 학습 과정에서 데이터의 특징을 자체적으로 추출하는데, 영상 기기나 질환 부위에 상관없이 일관된 특징 추출 모델을 가진다[16]. 이는 머신 러닝 보다 딥 러닝이 의료영상분석에서 더 많은 수효를 보이는 주요한 요인으로 여겨진다. 또한, 일관된 특징 추출 알고리즘을 사용하므로 이미 학습된 모델을 재사용하여, 다른 병변에 대해 추가로 학습시키는 전이 학습(transfer learning)이 가능한데, 일반적으로는 완전히 새로운 인공지능 모델을 사용해서 처음부터 학습시키는 것 보다는 이미 유의미한 판독 능력을 보이는 인공지능 모델에 전이 학습을 수행하면 학습의 속도가 빨라지거나 최종 알고리즘의 판독 성능이 더 뛰어난 것으로 여겨진다[17]. 다시 말해, 전이 학습은 이미 특정 병변에 대해 유의미한 판독 성능을 보이는 인공지능 모델이 의료영상의 경계선, 명도, 형상을 인식하는 기본적인 성능이 뛰어날 것이라는 기대에 근간을 둔다. 딥 러닝 기법도 정답 값인 레이블링 유무에 따라 지도 학습과 비지도 학습으로 나뉜다(Fig. 2).

III. 딥 러닝 기반 의료영상 분석의 원리

1. 합성곱 신경망

1) 합성곱 층

입력층의 영상 데이터가 첫 번째로 전달되는 층은 합성곱 층이다. 합성곱 층(Convolution Layer)은 합성곱 신경망

(Convolutional Neural Network; CNN)의 핵심적인 계층으로 입력 영상과 특징 커널 간의 유사도를 연산한다(Fig. 3). 특징 커널이란 학습 초기에는 난수 배열에 불과하지만, 학습이 진행되면서 판정의 대상이 되는 질환이 가장 많이 가지고 있는 영상 패턴으로 최적화되는 공간 필터로 이해할 수 있다. 합성곱 신경망에서는 특징 커널을 입력 영상 상에서 일정 간격으로 이동하며 합성곱 연산을 수행한다(Fig. 4). 윈도우와 특징 커널의 사이즈가 $n \times n$ 일 경우, 합성곱 연산은 아래의 식(1)으로 나타낼 수 있다[21].

$$z = x_{11}w_{11} + x_{22}w_{22} + \dots + x_{nn}w_{nn} \quad (1)$$

이때, x_{ij} 는 윈도우 내부의 입력 영상의 요소이고 w_{ij} 는 특징 커널의 요소이다. z 는 해당 영역의 특징 커널에 대한 유사도이다. 입력 영상과 특징 커널의 합성곱 결과인 유사도로 이루어진 행렬을 특징 지도(feature map)라 한다. 특징 지도는 특징 커널에 대한 입력 영상의 유사도를 나타내는 데이터이다. 특성 지도의 각 요소에 편향을 합성한 값을 가중 입력이라 하고, 이는 합성곱 층의 출력 데이터이다. Fig. 3의 예시는 숫자 2를 나타낸 6×6 입력 영상과 대각선 성분의 3×3 특징 커널의 합성곱 연산 과정이다. 출력 데이터의 유사도 분포를 확인하면, 입력 영상 중 특징 커널과 유사하게 대각선 패턴을 띄는 부분에서 높은 유사도를 보인다. 출력 데이터는 입력 영상과 비교하면 사이즈가 축소됨을 확인할 수 있는데, 일반적으로 딥 러닝에는 고도의 연산량이 요구되므로 합성곱 층을 통한 데이터 축소는 연산 속도 증가에 중요한 역할을 한다.

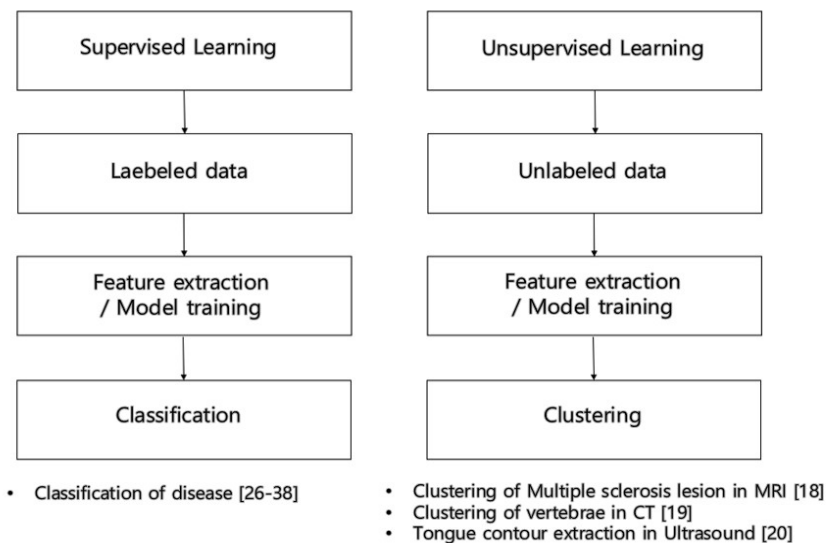


Fig. 2. Supervised learning and unsupervised learning

2) 활성화 층

활성화 층(Activation Layer)은 임계점 아래의 유사도 데이터를 정류하는 계층이다(별도의 계층이 아닌 합성곱 층의 출력 함수로 고려하는 경우도 있다). 즉, 특징 커널과의

유사도가 낮은 영역의 데이터를 학습 과정에서 배제하기 위함인데, 일반적으로 ReLu(rectified linear unit) 함수를 사용한다. 함수는 아래의 식(2)로 나타낼 수 있다[23].

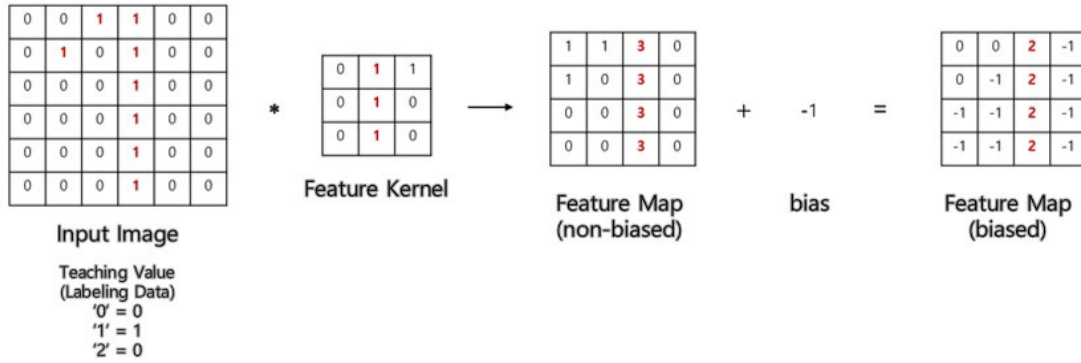


Fig. 3. Calculation process of the convolution layer

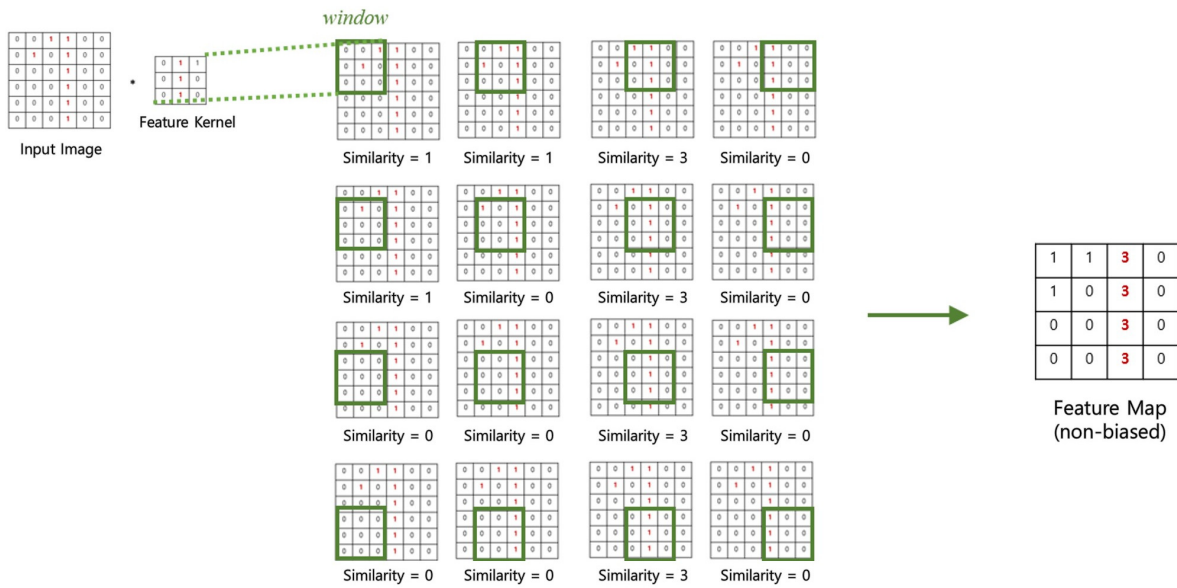


Fig. 4. Convolution between each window and the feature kernel

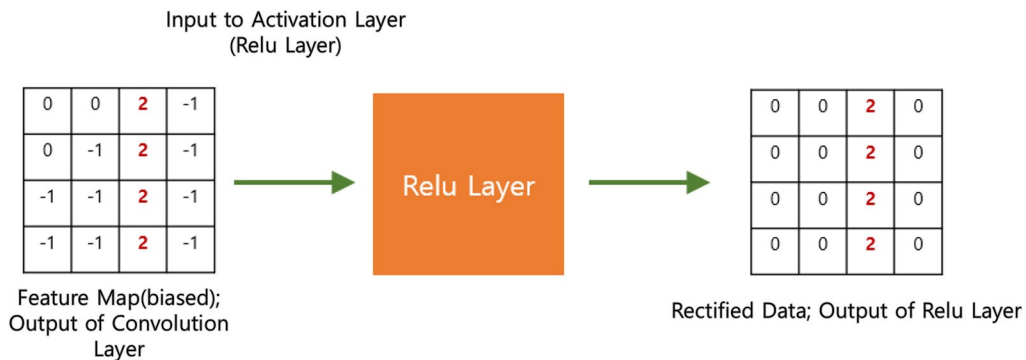


Fig. 5. Calculation process of the ReLu layer

$$a_{ij} = \text{ReLu}(z_{ij}) = \max(0, z_{ij}) \quad (2)$$

이때, z_{ij} 는 합성곱 층의 출력 데이터의 각 요소이고, a_{ij} 는 ReLu 함수의 출력값인 가중 입력이다. 입력값 z_{ij} 에 대해, $z_{ij} > 0$ 이면 출력의 기울기는 1이고, $z_{ij} < 0$ 이면 출력은 0이 된다(Fig. 5). 따라서, 특징 커널과 유의한 유사도를 나타내는 영역에 대해서만 학습을 수행할 수 있게 된다. 아래의 그림을 통해 활성화 층의 연산을 확인할 수 있다[21].

3) 풀링 층

풀링 층(Pooling Layer)에서는 합성곱 층과 활성화 층을 거치며 분류된 유의미한 유사도 데이터를 압축 계층이다. 막대한 연산량이 필요한 딥-러닝 학습모델에서 연산량을 줄이는 역할을 한다. 풀링은 압축 방법에 따라 최대 풀링(max pooling), 평균 풀링(mean pooling), 확률 풀링(stochastic pooling)으로 나뉜다[12].

- 최대 풀링: 임의의 영역($m \times n$) 내에 있는 요소 중 최댓값을 대푯값으로 출력하는 연산 방법(Fig. 6)
- 평균 풀링: 임의의 영역($m \times n$) 내에 있는 요소의 평균을 대푯값으로 출력하는 연산 방법
- 확률 풀링: 임의의 영역($m \times n$) 내에 있는 요소 중 하나를 확률에 근거하여 대푯값으로 설정하는 방법

의료영상 분석에는 최대 풀링이 흔히 사용되고 있다[24].

4) 출력층

출력층(Output Layer)은 합성곱 층, 활성화 층, 풀링 층을 거쳐 압축된 원본 영상의 유의미한 정보들을 하나의 판정 값으로 연산하는 계층이다(Fig. 7). n 의 분류를 수행할 수 있는 합성곱 신경망은 입력 영상에 대해 n 개의 신경망으로부터 출력된 n 개의 풀링 출력 데이터를 입력으로 받아들

인다. 각 출력 유닛은 모든 풀링 층으로부터 출력 값을 받는다. n 번째 출력 유닛의 연산은 아래의 식(3)으로 나타낼 수 있다[23].

$$z_n^o = w_{1-11}^{On} a_{11}^{P1} + w_{1-12}^{On} a_{12}^{P1} + \dots + w_{2-11}^{On} a_{11}^{P2} + w_{1-12}^{On} a_{12}^{P2} + \dots + w_{n-11}^{On} a_{11}^{Pn} + w_{n-12}^{On} a_{12}^{Pn} + b_n^o \quad (3)$$

이때, a_{ij}^{Pk} 은 k 번째 풀링 층으로부터 출력된 i 행 j 열의 요소이고, w_{k-ij}^{On} 는 n 번째 출력 유닛의 가중치이다. b_n^o 은 n 번째 출력 유닛의 편향이며, z_n^o 는 n 번째 출력 유닛의 출력 값이다. 이후 출력 값 z_n^o 에 활성화 함수를 적용하면 최종 출력 값 a_n^o 이 출력되며, 이를 해당 출력 유닛의 예측값이라고 한다. 이때 주로 사용되는 활성화 함수는 시그모이드 함수로 아래의 식(4)로 나타낸다[23].

$$a_n^o = a(z_n^o) = \frac{1}{1 + e^{-z_n^o}} \quad (4)$$

이후에 판정 값과 정답 값의 차이를 제곱 오차로 나타낸다. 정답 값은 사전에 영상에 레이블링 된 값으로, n 개의 판정이 가능한 분류기의 경우 n 개의 정답 값을 가진다. n 번째 판정이 정답일 경우, t_n 을 1, 그 외의 정답 값을 0으로 입력한다. m 개의 학습용 영상을 입력하여 n 개의 판정을 분류하는 신경망의 제곱 오차는 아래의 식(5)로 나타낼 수 있다[25].

$$C_k = \frac{1}{2} \{ (t_1 - a_1^o)^2 + (t_2 - a_2^o)^2 + \dots + (t_n - a_n^o)^2 \} \quad (5)$$

이때, C_k 는 학습용 입력 영상 중 k 번째 영상의 제곱 오차를 나타낸다($k=1, 2, 3, \dots, m$). t_n 은 각 영상에 레이블링 된 n 번째 판정에 대한 정답 값이다. a_n^o 은 n 번째 판정에 대

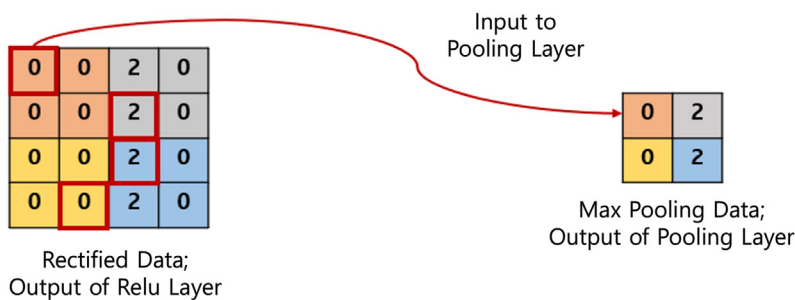


Fig. 6. Data compression via the pooling layer

한 예측값이다. 이는 예측값이 정답 값에 비해 큰 경우와 작은 경우의 차이를 모두 고려하기 위한 연산이며, 단일 영상에 대한 오차이다. 모든 학습 영상에 대한 오차의 총합은 비용함수 C_T 라 하고 아래의 식(6)으로 나타낸다[25].

$$C_T = C_1 + C_2 + C_3 + \dots + C_m \quad (6)$$

이때, C_m 는 m 입력 영상의 제곱 오차이고, C_T 는 비용함수를 나타낸다. 신경망을 이용한 학습 과정은 이 비용함수를 최소화하기 위해 파라미터(가중치, 편향, 필터)를 조정해나가는 과정으로 정의될 수 있으며, 이 비용함수가 최솟값을 나타내면 학습이 종료된다. 즉, 이 시점을 인간이 제공한 사전 지식을 습득하는 지점으로 여긴다.

2. 오차역전파법

비용함수를 최솟값이 되도록 파라미터를 정하기 위해서 제곱 오차와 각 파라미터를 함수 관계로 고려한다. 함수가 최솟값이 되기 위한 필요조건은 그 함수를 미분한 도함수가 0이어야 함을 응용한다. 하지만, 수많은 가중치, 편향이 함수와 복잡하게 얽혀 있는 신경망에서 연쇄적인 미분 연산을 하는 것은 비효율적인 일로 여겨진다. 이를 해결하기 위해 등장한 것이 오차역전파법(back propagation)이다[25].

오차역전파는 대규모 미분 계산을 수열의 점화식으로 대체한다. 수열의 점화식이란 유한수열 $\{a_n\}$ 이 있을 때, n 번째 항 a_n 과 인접한 $n+1$ 번째 항을 앞항과의 관계인 a_{n+1} 로 정의

하는 것을 의미한다. 이때, 사용한 관계식을 점화식이라고 한다. 이 점화식을 제공하는 변수를 유닛의 오차라고 한다. 유닛의 오차란 모든 층의 모든 값(유닛)에 대해서 정의할 수 있는데, 임의의 층 l 에서 j 번째 유닛의 가중치에 대한 유닛의 오차 δ_j^l 는 아래의 식(7)으로 나타낼 수 있다.

$$\delta_j^l = \frac{\partial C}{\partial z_j^l} \quad (7)$$

이때, C 는 단일 영상의 제곱 오차, z_j^l 는 임의의 층 l 의 j 번째 유닛의 가중치이며, δ_j^l 은 임의의 층 l 의 j 번째 유닛의 오차이다. 즉, 유닛의 오차는 제곱 오차에 대한 단일 가중치의 편미분이며, 임의의 가중치 z_j^l 가 제곱 오차에 미치는 변화율을 의미한다. 즉, 제곱 오차의 최솟값이 되는 조건에서 z_j^l 에 의한 제곱 오차의 변화율은 0이며, 이를 유닛의 오차 0으로 나타낸다. 이와 유사하게 임의의 층 l 의 j 번째 유닛의 b_j^l 의 편향 편향에 대한 유닛의 오차도 아래의 식(8)으로 나타낼 수 있다.

$$\delta_j^l = \frac{\partial C}{\partial b_j^l} \quad (8)$$

즉, 오차역전파법이란 출력층, 풀링 층, 활성화 층, 합성곱 층의 역순으로 각 유닛의 오차를 연산하는 기법이다. 상술하였듯이 각 유닛의 오차는 해당 가중치 혹은 편향이 제

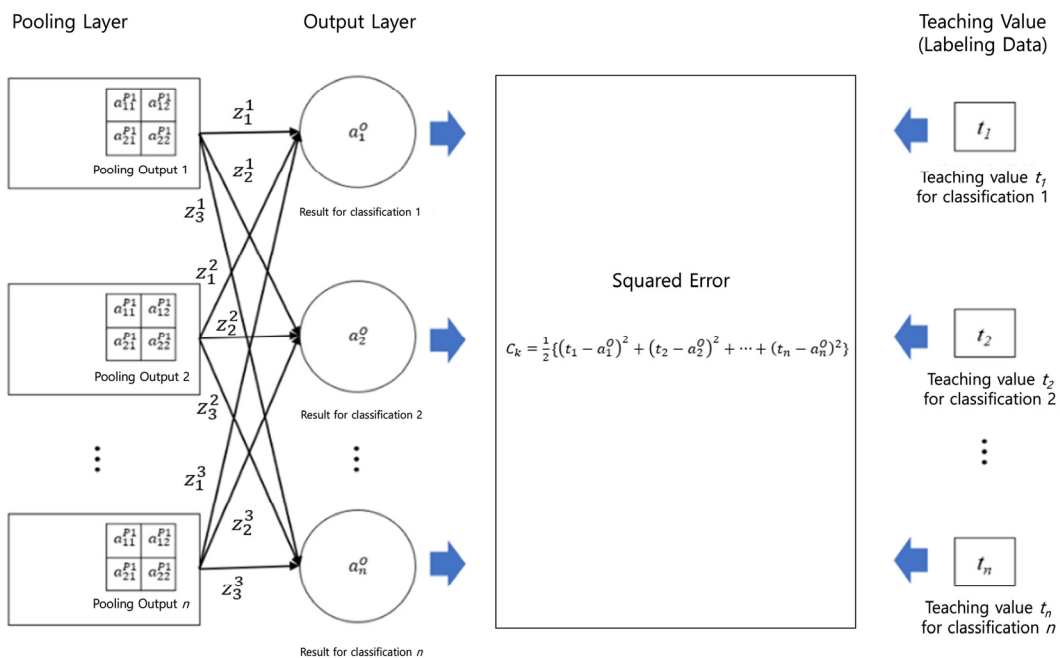


Fig. 7. Calculation process of the output layer

곱 오차에 미치는 순간 변화율을 의미한다. 편미분 값이 0이 되면 함수의 순간 변화량이 0이므로 최소값일 확률이 있으나, 변화량이 0이라 하더라도 극솟값이 아닌 경우가 더러 있다. 따라서, 수차례 반복 학습을 통해 모든 파라미터에 대해 최적화를 수행하는 것이다.

오차역전파 과정이 끝나면 n 개의 입력 영상에 대하여 모든 층의 모든 파라미터의 유닛의 오차 δ_j^l 가 계산된다. n 개의 입력 영상에 대한 임의의 유닛 오차 δ_k 를 통해서 다음 회차에 사용될 파라미터를 갱신하는 과정은 아래의 식(9), (10)을 통해서 확인할 수 있다.

$$\delta_{1st\ training} = \sum_{k=1}^n \delta_k \quad (9)$$

$$P_{2nd\ training} = P_{1st\ training} - \delta_{1st\ training} * \eta \quad (10)$$

여기서, δ_k 는 전 회차 학습의 k 번째 입력 영상의 임의의 층 l 에서 j 번째 유닛의 파라미터의 유닛의 오차, $\delta_{1st\ training}$ 은 n 개의 입력 영상의 같은 층 l 에서 j 번째 유닛의 오차 합산값, $P_{1st\ training}$ 은 l 층의 j 번째 파라미터의 전 회차 학습에 사용되었던 값, η 는 학습률, $P_{2nd\ training}$ 은 다음 회차 학습에서 l 층의 j 번째로 사용될 파라미터 값이다. 즉, 1개의 유닛을 갱신 값은 모든 영상에 대해서 그 유닛 값을 합산하고 전 회차 가중치에 감산하는 것과 같다. 하지만, 학습률이란 파라미터 값이 너무 큰 폭으로 갱신되어 그사이에 존재하는 값이 적용되지 못하는 것을 방지하기 위한 개념으로, 파라미터가 갱신되는 정도를 작게 하여 정밀한 최적화를 돕는 값이다. 학습률은 상황에 따라 가변적이지만 0.2를 사용하는 것이 일반적이다. 위의 과정을 통해서 모든 파라미터의 갱신이 완료되었다면, 갱신된 파라미터들로 다음 회차의 학습을 수행하고 비용함수를 확인하면 된다. 수차례의 파라미터 갱신과 학습을 통해 비용함수가 최소값을 보이면 학습을 중단한다. 이때, 대부분 유닛의 오차는 0이다. 이 상황을 학습이 종료된 것으로 여기는 것이 일반적이다.

3. 실사용 예시

딥 러닝을 이용해서 흉부 X선 영상의 폐렴과 폐암 여부를 진단하는 경우, 그 학습 및 판독과정은 다음과 같다.

1) 학습 과정

학습 과정의 첫 단계로 의사의 진단 결과를 딥 러닝에게 학습시켜야 한다. 의료영상에 학습의 정답을 표기하는 것을

레이블링(labeling)이라고 하는데, 대부분 소프트웨어는 그 편의성 때문에 폴더의 이름을 판정 값으로 불러올 수 있는 폴더 레이블링을 지원한다. 예를 들어, 흉부 X선 영상을 통해 폐렴과 폐암을 진단하는 알고리즘을 학습시킬 경우, 폐렴으로 진단된 영상을 "폐렴" 이름의 폴더에 넣고, 폐암의 영상을 "폐암" 이름의 폴더에 넣으면 소프트웨어상에서 "폐렴"과 "폐암" 두 가지 판정 선택지가 생성된다. 이때, "폐렴" 폴더에 있는 영상들은 폐암=0, 폐렴=1, "폐암" 폴더에 있는 영상들은 폐암=1, 폐렴=0의 정답 값이 표시(labeling)되는 형태가 일반적이다. 이때, 1과 0이란 각 영상이 가지고 있는 특징에 대한 해당 질환 판정의 일치율이다. 즉, 각 질환 판정에 대해 100%의 일치율을 보이는 레이블링 영상들은 알고리즘의 입장에서 해당 판정을 내리기 위한 영상 특징의 정답으로 설명될 수 있다. 이 정답 값들은 상술한 것과 같이 비용함수를 계산하기 위해 사용된다. 수차례의 학습과 파라미터 갱신을 걸쳐 비용함수가 최소값을 나타내면 알고리즘이 폐암과 폐렴을 분류할 적절한 성능을 띄며, 이를 "폐암, 폐렴 분류 모델"로 고려한다.

2) 판독과정

상술한 과정을 거쳐 의사의 사전 지식이 충분히 학습된 "폐암, 폐렴 분류 모델"에 흉부 X선 영상을 입력할 경우, 학습을 통해 최적화된 폐암을 분류해낼 수 있는 특징 커널과 폐렴을 분류해낼 수 있는 특징 커널에 대해 합성곱 연산을 수행한 응답인 특징 지도가 계산된다. 이후 활성화 함수 계산과 풀링 층, 출력층을 거쳐 정보가 압축되며 각 분류에 대한 판정 값이 출력된다.

IV. 연구 동향

현대 AI 의료영상 분석 기술은 일반촬영술을 시작으로 전산화단층촬영, 자기공명영상, 초음파촬영, 투시 촬영 등의 다양한 영상 장비에 적용되고 있으며, 다양한 질환을 대상으로 유의미한 판독 결과가 보고되고 있다[26-38].

1. 흉부

일반촬영을 통한 흉부 질환 분류는 가장 유의한 정확도를 보이는 분야이다. CNN의 일종인 CheXNeXt 네트워크가 가장 널리 이용되고 있으며, 무기폐(atelectasis), 심장비대(cardiomegaly), 폐경화(consolidation), 부종(edema), 흉막삼출(effusion, emphysema), 섬유증(fibrosis), 탈장(hernia),

침윤(infiltration), 종양(mass), 결절(nodule), 횡경막 비후화(pleural-thickening), 폐렴(pneumonia), 공기기슴증(pneumothorax) 등 AP 영상만으로도 분류할 수 있는 질환의 종류가 다양하다. Rajpurkar 등[26]이 14종의 흉부 질환에 대해 CheXNeXt 알고리즘과 실제 의사의 정확도를 비교한 결과, 결핵 및 폐암 등을 포함한 11종의 주요 질환에 대해서 실제 의사와 유사한 곡선하면적(area under the curve, AUC)을 보였으며, 그 중 무기폐 소견에 대해서는 0.862 AUC를 보여 실제 의사의 0.808와 비교하여 유의하게 높은 정확도를 나타냈다. 하지만, 이러한 결과는 영상에서 비대칭성을 보이거나 국소 부에 관찰되어 비교적 판단의 근거가 명확한 질환에 한정된다. 그 외의 폐기종과 열공 탈장 등의 병변에서는 낮은 정확도를 보였는데, 폐기종의 경우 넓은 범위에 걸쳐 대칭적으로 병변이 관찰되기 때문에 병변만의 특징을 학습시키기에 어려움이 있고, 열공 탈장은 여타 질환에 비해 상대적으로 드문 증례로서 학습 데이터의 수가 적은 것이 이유인 것으로 추정된다. 더욱이, 상기한 유의미한 판독 결과는 양쪽 흉부의 정보를 모두 포함하는 AP(anteroposterior) 혹은 PA(posteroanterior) 영상을 토대로 이루어진 것 또한 현재 기술의 한계이다.

2. 심장

심장 질환 분야에서는 관상동맥질환을 진단하는 관상동맥 전산화 단층 혈관조영술(coronary artery computed tomography angiography; CCTA)을 통한 관상동맥 석회화 수치(coronary artery calcium scoring; CACS)를 자동으로 평가하는 기술이 탁월한 성능을 보인다. 관상동맥석회화 수치(coronary artery calcium score; CAC)는 관상동맥의 죽상경화증을 평가하는 지표이며, 심전도 동기 하의 비 조영증강 전산화 단층촬영영상으로 평가한다. 석회화 병변이 명도 등에서 이상값을 보인다는 특성을 활용한다. Shadmi 등[27]은 합성곱신경망의 일종인 U-Net과 DenseNet을 이용해서 조영증강 없는 전산화단층촬영영상에서 관상동맥 석회화 수치를 예측하였다. 그 결과, 실제 의사의 레이블링에 대해 0.98의 높은 피어슨 상관 계수를 보였고, 두 관측치의 일치도를 확인하기 위해 블랜드-앨트먼(Bland-Altman plot) 분석을 시행한 결과 95% 일치한계에서 0.4의 낮은 평균오차(bias)를 나타냈다. 하지만, 저자에 따르면 승모판(mitral valve)에서의 거짓 양성 판정과 대동맥과 관상동맥이 밀접한 경우 등에서의 오판독을 현재 기술의 한계점으로 보고하고 있다.

3. 뇌

1) 뇌종양

뇌종양 분야에서는 병변을 자동으로 분할하는(segmentation) 딥 러닝 기반 기법이 활발하게 연구되고 있다. 가장 널리 사용되고 있는 데이터 셋은 BRATS(Brain Tumor Image Segmentation Benchmark)으로, 274장의 뇌종양 자기공명영상에 대해 방사선 전문의의 뇌종양 분할 정답 값이 입력되어 있는 데이터 셋이다[28]. 대부분의 연구에서는 이 BRATS 데이터 셋에 제시된 뇌종양 분할의 정답 값과의 유사도를 뇌종양 학습모델의 성능지표로 보고하고 있다. 이 유사도는 Dice score로 나타내며 아래의 식(11)으로 나타낸다.

$$Dice\ score = \frac{2 \times TP}{(TP + FP) + (TP + FN)} \quad (11)$$

BRATS 데이터 셋의 정답 값에 대한 의사 그룹의 Dice score 지표로는 Menze 등[28]이 제시한 의사 그룹의 Dice score를 일반적으로 사용하는데, whole 종양에 대해 0.85, core 종양에 대해 0.75, active 종양에 대해 0.74를 나타낸다. Urban 등[29]은 합성곱신경망 기반의 딥 러닝 학습모델을 이용하여 BRATS 데이터 셋의 영상의 뇌종양 병변을 분할한 결과가 whole 종양에 대해 0.87, core 종양에 대해 0.77, active 종양에 대해 0.73을 나타내어 core 종양과 active 종양에 대해 의사 그룹보다 낮은 정확도를 나타냄을 보고하였다. 이와 유사하게 Zikic 등[30]은 합성곱신경망 기반의 딥 러닝 학습 모델을 이용하여 BRATS 데이터 셋의 영상의 뇌종양 병변을 분할한 결과가 whole 종양에 대해 0.84, core 종양에 대해 0.74, active 종양에 대해 0.69를 나타내어 core 종양과 active 종양에 대해 의사 그룹보다 낮은 정확도를 나타냄을 보고하였다.

2) 뇌졸중

뇌졸중은 뇌에 혈액을 공급하는 혈관이 막히거나 터져서 발생하는 급성 질환으로, 발생 후 시간이 지날수록 예후가 좋지 못하므로, 빠른 진단과 치료가 중요한 질환이다. 즉, 뇌졸중 진단 분야에서 인공지능 판독 기술의 이점은 신속성이다. 주로 전산화단층촬영 혈관조영술(computed tomography angiography; CTA)에 합성곱신경망 기반의 딥 러닝 모델로 판독하는 기술이 연구되고 있다. Omen 등[31]은 합성곱신경망의 일종인 DeepMedic 소프트웨어를 이용하여 CTA로부터 허혈성 뇌졸중 병변 분류를 시도하였다. 사전 학습을 위해 전문의가 수작업으로 뇌졸중 병변에 ROI를 표기한 분할 영상이 사용되었다. 방사선 전문의에 의한 수작업으로

분할된 병변과 비교한 결과, DeepMedic 소프트웨어는 0.93의 민감도와 0.82의 특이도를 나타내어 유의한 수준의 판독 능력을 나타냈다. 현재 다양한 뇌졸중 촬영 알고리즘을 바탕으로 뇌졸중의 진단 속도 및 영상의 질이 향상되어, 성공적인 뇌졸중 진단과 치료가 이루어지고 있지만[32-34], 향후 발전하는 AI 기술은 더욱 향상된 의료 서비스를 제공할 수 있을 것이라 기대된다.

3) 알츠하이머병

알츠하이머병은 자기공명영상 혹은 PET 영상을 통한 자동화 진단, 경도 인지 장애(mild cognitive impairment; MCI) 환자의 알츠하이머 전환을 예측하는 연구가 진행되고 있다. 알츠하이머병 초기에는 뇌의 국소부위에서 미세한 해부학적 변화가 있으므로 전체 뇌 영상의 정보를 이용하여 알츠하이머를 초기에 판독하는 것은 다소 비효율적이라는 지적이 있다[35]. Liu 등[35]은 상기한 문제를 극복하기 위해 뇌의 전체가 아닌 국소 랜드마크에서 특징을 추출하여 알츠하이머를 초기에 진단하는 기법을 제안하였다. 랜드마크는 정상 소견 환자와 알츠하이머 환자의 데이터를 비교하여 해부학적 변화를 나타낸 50개의 지점을 선택하여 판독을 진행하였다. 이 기법은 알츠하이머 진단에 0.96의 높은 AUC를 보이는 것으로 나타났다. Choi 등[36]은 18F-FDG PET 영상에 합성곱신경망을 적용하여 경도 인지 장애 환자의 알츠하이머 전환을 84.2%의 정확도로 예측하는 결과를 보고하기도 하였다.

4. 유방

유방 분야에서는 대부분 유방암을 자동으로 진단하는 연구가 진행되고 있다. Huynh 등[37]은 합성곱신경망 기반의 딥 러닝 모델을 전이 학습시켜 디지털 유방촬영 영상에서 자동 유방암 진단을 하였고, 0.86의 유의한 AUC를 보임을 보고하였다. 또한, 2020년 구글 헬스 연구소의 연구 결과에 따르면[38], AI와 실제 의사의 유방암 판독을 비교한 결과, AI의 허위 양성(false positive) 비율이 방사선 전문의 진단 결과보다 5.7% 적었고, 허위 음성(false negative) 비율은 9.4%가 적었다. 이러한 결과는 AI가 인간 의사의 진단 효율을 높이기 위한 보조 도구로 활용될 수 있음을 시사하는데, 허위 양성에 의한 불필요한 조직 생검(biopsy)이나 재촬영으로 인한 피폭의 부담을 줄일 수 있고, 허위 음성에 의한 오진을 줄일 수 있다.

V. 고찰

1. 보조 진단 기구로서의 딥 러닝

상기한 여러 가지 사례들은 딥 러닝 기반의 의료영상 판독 기술이 유의미한 성능을 나타냄을 뒷받침하며, 이에 많은 연구진이 딥 러닝 기반 의료영상 판독 기술이 의사를 보조할 수 있는 보조 진단 기구(computer aided diagnosis; CAD)로 사용될 수 있음을 강조했다[38-42]. 다시 말해, 의사의 임상적 지식과 경험, 환자에 대한 이해와 딥 러닝의 빅데이터 학습 결과를 모두 반영함으로써 진단 효율을 극대화하는 것으로, 딥 러닝을 진단의 주체가 아닌 보조 도구의 성격으로 고려하는 것이다. 이러한 전망에는 딥 러닝의 빠른 처리 속도가 긍정적인 기여를 하는데, Rajpurkar 등[26]의 실험결과에 따르면 방사선 전문의 그룹이 420장의 의료영상을 판독하기 위해 평균 240분이 걸렸지만, 딥 러닝은 동일 조건에서 1.5분이 소요된 것으로 나타났다. 본래 의료영상 분석은 노동집약적 업무로서 판독이 난해한 질환의 경우에는 수 시간이 걸리는 일도 있어, AI의 보조는 의료진에게 가해지는 노동 부담을 상당수 덜어줄 것으로 전망된다. 또한, Mckinney 등[38]의 유방암 판독 실험결과에 따르면 딥 러닝이 허위 음성에 의한 오진이 방사선 전문의 그룹에 비해 9.4% 낮게 나타났다. 대부분 질환이 대처 시점이 늦어질수록 예후가 나빠짐을 고려하면, 딥 러닝의 낮은 허위 음성 수치는 인간의 실수에 의한 오진을 줄이는 데 긍정적인 영향을 미칠 것이다. 또한, 영상의학 인력이 부족한 상황에서의 활용 가능성을 크게 평가받고 있는데, Dross 등[43]의 2014년 조사결과에 따르면, 일부 국가에서는 인구 100만 명당 영상의학 의사가 10명 이하인 경우도 있다. 국내의 경우, 응급 영상의학 인력의 공백이 지적됐는데, 대한영상의학회에 따르면 2017년 기준 영상의학 전문의 약 3700명 중 응급실에서 영상 판독을 전담하는 의사는 0.27%에 그치는 10여 명으로 나타났다[44]. 이러한 인력 부족 현상은 판독 지연 및 업무 과중으로 인한 오진으로 이어져 환자 케어에 악영향을 미칠 수 있으므로, 딥 러닝이 가까운 미래에 이를 보완하는 솔루션으로 기여할 수 있다.

2. 데이터 셋

현대 딥 러닝 기반 의료영상 분석 기법의 빠른 성장세는 다양한 데이터 셋의 공유에 기인한다. 데이터 셋이란 신뢰할 수 있는 전문가 집단에 의해 레이블링이 수행된 대규모 의료영상 데이터를 의미하며, 이 중 일부를 사전 학습에 사

용하고 나머지를 판독 성능 검증에 사용하는 것이 일반적이다. 의료영상에서의 데이터 셋은 빅 데이터의 5대 요소인 양(volume), 다양성(variety), 처리속도(velocity), 신뢰도(veracity), 수치(value)의 특성을 지니고 있다. 의료영상 데이터 셋은 다양한 공급자로부터 공유되고 있는데, 대표적으로 흉부 일반촬영상을 위한 Chest X-ray[45], 자기공명 촬영 영상의 알츠하이머병 영상을 포함하는 ADNI-1[46], MIRIAD[47], 각종 유방 질환의 영상을 포함하는 INbreast 등이 있다[48]. 하지만, 여전히 일부 질환에 대해서는 주석이 달린 의료영상 빅 데이터를 구하는 것은 순조롭지 않은 것으로 여겨진다. 의료영상을 판독해서 주석을 다는 것은 전문 인력이 필요하고 많은 시간이 소모되고, 의사 개인에 따라 상이한 소견을 가질 수 있다. 이러한 제한점은 희귀한 증례를 다루는 데이터 셋을 확보하기 어렵게 만든다[49]. 또한, 데이터 셋에 관한 윤리적 문제도 짚고 넘어가야 하는데, 의료영상은 환자의 사적 정보이므로 이를 다루는 인력은 환자의 비밀보장, 사적정보 보호 등의 윤리적 책임이 포함된다. 따라서, 데이터 셋의 익명화와 제한된 접근이 중요할 것으로 여겨지나, 이는 데이터 셋의 볼륨을 줄일 가능성이 있고, 데이터 접근성을 줄인다는 단점도 수반하므로 적절한 해결책이 마련되어야 할 것이다.

3. 연구전망

의료영상 분야에서의 딥 러닝 기술은 초입 국면으로, 구글 딥마인드, IBM 왓슨 등의 대형 기업의 연구소를 중심으로 연구 생태가 조성되고 있다. 특히, IBM은 의료진과 협업하여 실제 의료현장에 접목할 수 있는 어플리케이션을 연구하고 있다. 구글 딥마인드는 영국의 국민 건강 서비스(National Health Service, NHS)와의 협업을 통해 실명 진단을 위한 홍채 분석 등 다양한 의료 서비스 어플리케이션 등의 연구를 수행하고 있다. GE Healthcare 와 Siemens 등 의료기기 회사들도 딥 러닝 기술에 투자를 하고 있는데, GE healthcare 사는 보스턴 아동 병원과의 협업을 통해 소아 뇌질환을 예측하는 스마트 영상화 기술을 개발하고 있고, UC San Francisco와의 협업을 통해 정상 소견과 의사의 더 많은 주의를 필요로 하는 소견으로 구분할 수 있는 알고리즘을 개발 중이다[49].

최근에는 의료영상을 분할하거나 판독하는 것 이외에 딥 러닝을 이용한 화질 개선 기법도 주목을 받고 있는데, 일례로 저선량 CT 영상의 화질을 딥 러닝으로 개선하는 기법[50, 51], 관류자기공명영상의 화질을 딥 러닝으로 개선하

는 기법[52], 저선량 PET의 영상 화질을 개선하는 기법 등이 있다[53]. 특히, CT, PET 등 방사선 피폭을 수반하는 영상화 기기에서 저선량 영상의 화질을 개선함으로써, 영상의 진단적 정보는 유지하면서도 환자의 피폭선량을 줄여 환자 케어의 수준을 높일 수 있을 것으로 기대된다.

Razzak 등[49]은 희귀한 증례에 대한 데이터 셋을 확보하기 어려움을 지적하며, 기존의 레이블링에 기반을 둔 지도학습(supervised learning) 딥 러닝 연구 영역이 데이터의 내적 관계만을 고려하는 비지도 학습 기법(unsupervised learning) 혹은 레이블링 데이터와 레이블링이 없는 데이터를 모두 학습에 사용하는 준 지도 학습(semi-supervised learning) 체제로 전환되어야 함을 제언했다. 또한, 의료영상에서 딥 러닝 분야가 지속해서 발전하기 위해서는 은닉층에 대한 설명력을 높여야 함을 시사했는데, 일부 의료진이 딥 러닝의 신뢰성에 의구심을 표하는 것이 이러한 설명력의 부재에 근거한다는 주장이다. 이는 성능과 별개로 인식에 관한 문제로 여겨지는데, 이를 극복하기 위해서는 의료영상 분야 외에도 자율운전, 로봇공학 등 실생활의 중요 분야에서 딥 러닝이 신뢰할 만한 성능을 보이고, 은닉층에 대한 설명력을 개선해야 함을 제언했다.

상술한 합성곱 신경망 이외에도 생성적 적대 신경망(Generative Adversarial Network; GAN)의 등장으로 관련 연구 분야가 더욱 확대되고 있다[54]. 생성적 적대 신경망이란 학습된 데이터를 바탕으로 원본 영상과 유사한 특징을 지닌 인공 영상 정보를 생성(generative)해낼 수 있는 신경망이다. 적대(Adversarial)라는 용어는 이 신경망이 생성자(generator)와 식별자(discriminator)로 나뉘어 적대적 학습을 수행하는 것을 의미하는데, 적대적 학습이란 생성자가 인공 데이터를 만들어내어 식별자를 속이고, 식별자는 인공 데이터와 실제 데이터를 구분하는 일련의 학습 과정을 의미한다. 이들을 동시에 학습시키는 결과로 생성자는 실제 데이터와 구분할 수 없는 인공 데이터를 생성하고, 식별자는 뛰어난 식별 성능을 가지게 된다. 의료영상 분석에는 주로 화질 개선, 해상도 향상 등의 목적으로 전산화단층촬영과 자기공명영상 분야에서 다양한 연구가 이루어지고 있고[55-58], 단일 환자의 자기공명영상으로부터 그에 대응하는 전산화 단층 영상을 생성해내는 기법 등이 연구되고 있다[59]. 생성적 적대 신경망은 기존의 신경망에 비해 최적화가 매우 어렵다는 한계가 있지만, 단순히 수동적인 분류기의 역할을 벗어나 인공 정보를 능동적으로 만들어 낼 수 있다는 점에서 AI 연구의 새 지평을 열게 될 것이라 여겨진다.

4. 제한점

현재 기술 수준에서는 상술한 것과 같이 합성곱 신경망 기반의 딥 러닝 기법은 다양한 질환에 대해서 유의미한 결과를 나타내고 있다. 하지만, 딥 러닝 기술은 몇 가지 제한점을 가지고 있다. 첫 번째는 특정 질환의 병변 특징을 잘 나타내는 명료한(clear) 영상이 대규모로 필요하다는 것인데, 이는 레이블링도 수반해야 하므로 적절한 데이터 셋의 확보가 관건이다. 두 번째는 딥 러닝이 오로지 영상의 정보만을 반영하여 환자 개인에 대한 사전 정보 등을 고려할 수 없는 것도 한계점으로 지적된다. 세 번째로 딥 러닝은 필터, 가중치, 편향, 학습률 등 수많은 파라미터에 대해서 매우 민감한 기법이므로 이미 학습된 모델의 일부 바이어스만을 수정하는 것이 매우 어렵다는 단점이 있다. 이를 극복하기 위해서는 학습 데이터의 양을 늘려야 하지만 과적합(overfitting) 될 수 있다는 모순도 존재한다. 그리고 상술한 설명력의 부재는 딥 러닝이 극복해야 할 가장 중요한 사안이다.

VI. 결 론

본 연구에서는 딥 러닝 기반의 의료영상 판독 기술에 대해서 개념과 원리를 소개하였고, 문헌 고찰을 통해 관련 연구 현황을 논의하였다. 딥 러닝은 컴퓨터 연산 성능과 빅 데이터 공유의 증가로 빠르게 발전하고 있는 분야이며, 질환의 분류, 분할 등을 자동으로 수행할 수 있는 기법이다. 현재 일반촬영, CT, MRI, PET 등 다양한 영상화 기법에 적용되고 있으며, 다수의 질환에 대해서 인간 의사의 진단 정확도와 비교하여 유의미한 진단 정확도를 보이는 연구 결과가 다수 보고되고 있다. 하지만, 환자의 이력 등을 반영하지 않고 영상의 정보만을 토대로 판독을 수행하는 원리적 한계가 있다. 따라서, 이를 인간 의사를 대체할 진단의 주체로 여기기에는 어려움이 있고, 의사의 진단을 보조하는 진단 보조 도구(computer aided diagnosis; CAD)의 성격으로 보는 것이 옳다. 즉, 신속한 처리 속도와 육안으로 확인할 수 없는 미시적 수준에서의 병변 식별 능력 등 딥 러닝의 가지는 장점이 의사를 보조하여 오진을 줄일 수 있으며, 이를 통해 환자와 의사에게 많은 이익을 제공할 수 있을 것으로 기대된다.

REFERENCES

- [1] Harman M. The role of artificial intelligence in software engineering. In: 2012 First International Workshop on Realizing AI Synergies in Software Engineering (RAISE). IEEE, 2012 Jun:1-6.
- [2] Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *Eur Radiol Exp*. 2018 Oct; 2(1):35.
- [3] Trinder JC, Wang Y, Sowmya A, Palhang M. Artificial intelligence in 3-D feature extraction. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, 2nd ed. Basel: Birkhäuser; 1997.
- [4] Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng*. 2017 Jun; 19:221-48.
- [5] Rav D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, et al. Deep learning for health informatics. *IEEE J Biomed Health Inform*. 2017 Jan; 21(1):4-21.
- [6] LeCun Y, Boser B, Denker JS, Henderson D. Backpropagation applied to handwritten zip code recognition. *Neural Comput*. 1989;1(4):541-51.
- [7] Neyshabur B, Tomioka R, Srebro N. In search of the real inductive bias: On the role of implicit regularization in deep learning. *arXiv preprint arXiv:1412.6614*; 2014.
- [8] Yu K, Xu W, Gong Y. Deep learning with kernel regularization for visual recognition. *Adv Neural Inf Process Syst*. 2009:1889-96.
- [9] Kukačka J, Golkov V, Cremers D. Regularization for deep learning: A taxonomy. *arXiv preprint arXiv:1710.10686*; 2017.
- [10] Yoshida Y, Miyato T. Spectral norm regularization for improving the generalizability of deep learning. *arXiv preprint arXiv:1705.10941*; 2017.
- [11] Russell S, Norvig P. *Artificial intelligence a modern approach*, 3rd ed. New Jersey: Prentice Hall; 2009.
- [12] Mitchell TM. *Machine learning*. New York: McGraw-Hill; 1997.

- [13] Vial A, Stirling D, Field M, Ros M, Ritz CH, Carolan, MG, et al. The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: A review. *Translational Cancer Res.* 2018;7(3):803-16.
- [14] Ang JC, Mirzal A, Haron H, Hamed HNA. Supervised, unsupervised, and semi-supervised feature selection: A review on gene selection. *IEEE/ACM Trans Comput Biol Bioinform.* 2015;13(5):971-89.
- [15] Lin YZ, Nie ZH, Ma HW. Structural damage detection with automatic feature extraction through deep learning. *Computer-Aided Civ Inf Eng.* 2017;32(12):1025-46.
- [16] Yang J, Zhao YQ, Chan JCW. Learning and transferring deep joint spectral-spatial features for hyperspectral classification. *IEEE Trans Geosci Remote Sens.* 2017;55(8):4729-42.
- [17] Weiss KR, Khoshgoftaar TM. Comparing transfer learning and traditional learning under domain class imbalance. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2017 Dec:337-43.
- [18] Yoo Y, Brosch T, Traboulsee A, Li DK, Tam R. Deep learning of image features from unlabeled data for multiple sclerosis lesion segmentation. In *International Workshop on Machine Learning in Medical Imaging*. 2014 Sep:117-24.
- [19] Cai Y, Landis M, Laidley DT, Kornecki A, Lum A, Li S. Multi-modal vertebrae recognition using transformed deep convolution network. *Comput Med Imaging Graph.* 2016 Jul;51:11-9.
- [20] Jaumard-Hakoun A, Xu K, Roussel-Ragot P, Dreyfus G, Denby, B. Tongue contour extraction from ultrasound images based on deep neural network. *arXiv preprint arXiv:1605.05912*; 2016.
- [21] Saito K. *Deep learning starting from the bottom*. Seoul: Hanbit Media; 2017.
- [22] Kim H, Nam H, Jung W, Lee J. Performance analysis of CNN frameworks for GPUs. In *2017 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*. 2017 Apr:55-64.
- [23] Agarap AF. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*; 2018.
- [24] Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB. Deep learning in medical imaging: General overview. *Korean J Radiol.* 2017 Aug;18(4):570-84.
- [25] Wakui Y, Wakui S. *Math to understand deep learning*. Seoul: Hanbit Media; 2018.
- [26] Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Medicine.* 2018;15(11):e1002686.
- [27] Shadmi R, Mazo V, Bregman-Amitai O, Elnekave E. Fully-convolutional deep-learning based system for coronary calcium score prediction from non-contrast chest CT. In *2018 IEEE 15th International Symposium on Biomedical Imaging*. 2018 Apr:24-8.
- [28] Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging.* 2014;34(10):1993-2024.
- [29] Urban G, Bendszus M, Hamprecht F, Kleesiek J. Multi-modal brain tumor segmentation using deep convolutional neural networks. *MICCAI BraTS (brain tumor segmentation) challenge*. Proceedings, Winning Contribution. 2014:31-5.
- [30] Zikic D, Ioannou Y, Brown M, Criminisi A. Segmentation of brain tumor tissues with convolutional neural networks. *Proceedings MICCAI-BRATS*. 2014:36-9.
- [31] Öman O, Mäkelä T, Salli E, Savolainen S, Kangasniemi M. 3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke. *Eur Radiol Exp.* 2019;3(1):8.
- [32] Yang W, Hong JY, Kim JY, Paik SH, Lee SH, Park JS, et al. A novel singular value decomposition-based denoising method in 4-dimensional computed tomography of the brain in stroke patients with statistical evaluation. *Sensors.* 2020;20:3063.
- [33] Park HY, Pyeon D, Kim DH, Jung Y. Dynamic computed tomography based on spatio-temporal analysis in acute stroke: Preliminary study. *J Radiol Sci Technol.* 2016;39:543-7.

- [34] Kim D, Jung Y. Simulation study for feature identification of dynamic medical image reconstruction technique based on singular value decomposition. *J Radiol Sci Technol*. 2019;42:119–30.
- [35] Liu M, Zhang J, Adeli E, Shen D. Landmark-based deep multi-instance learning for brain disease diagnosis. *Med Image Anal*. 2018;43:157–68.
- [36] Choi H, Jin KH. Alzheimer's Disease Neuroimaging Initiative. Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. *Behav Brain Res*. 2018;344:103–9.
- [37] Huynh BQ, Li H, Giger ML. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *J Med Imaging*. 2016;3(3):034501.
- [38] McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020;577(7788):89–94.
- [39] Christe A, Peters AA, Drakopoulos D, Heverhagen JT, Geiser T, Stathopoulou T, et al. Computer-aided diagnosis of pulmonary fibrosis using deep learning and CT images. *Invest Radiol*. 2019;54(10):627.
- [40] Hamidinekoo A, Denton E, Rampun A, Honnor K, Zwiggelaar R. Deep learning in mammography and breast histology, an overview and future trends. *Med Image Anal*. 2018;47:45–67.
- [41] Cheng JZ, Ni D, Chou YH, Qin J, Tiu CM, Chang YC. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. *Sci Rep*. 2016;6(1):1–13.
- [42] Hua KL, Hsu CH, Hidayati SC, Cheng WH, Chen YJ. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *Onco Targets Ther*. 2015;8.
- [43] Dross PE, Sumner R, Tysowsky M, Aujero MP. International radiology for developing countries: The delivery of medical imaging services to a Honduran radiologic scarce zone. *J Am Coll Radiol*. 2014;11(12):1173–7.
- [44] Kwon DI. Only about 10 people in the emergency room nationwide who specialize in imaging medicine to read CT · MRI. *Hankook Ilbo*, 2017.11.20.
- [45] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. Chest X-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*. 2017:2097–106.
- [46] Jack Jr CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging*. 2008;27(4):685–91.
- [47] Malone IB, Cash D, Ridgway GR, MacManus DG, Ourselin S, Fox NC, et al. MIRIAD—Public release of a multiple time point Alzheimer's MR imaging dataset. *NeuroImage*. 2013;70:33–6.
- [48] Moreira IC, Amaral I, Domingues I, Cardoso A, Cardoso MJ, Cardoso JS. Inbreast: Toward a full-field digital mammographic database. *Acad Radiol*. 2012;19(2):236–48.
- [49] Razzak MI, Naz S, Zaib A. Deep learning for medical image processing: Overview, challenges and the future. In *Classification in BioApps*. 2018:323–50.
- [50] Ding Q, Chen G, Zhang X, Huang Q, Ji H, Gao H. Low-dose CT with deep learning regularization via proximal forward backward splitting. *Phys Med*. 2020.
- [51] Shin YJ, Chang W, Ye JC, Kang E, Oh DY, Lee YJ. Low-dose abdominal CT using a deep learning-based denoising algorithm: A comparison with CT reconstructed with filtered back projection or iterative reconstruction algorithm. *Korean J Radiol*. 2020;21(3):356–64.
- [52] Xie D, Li Y, Yang H, Bai L, Wang T, Zhou F. Denoising arterial spin labeling perfusion MRI with deep machine learning. *Magn Reson Imaging*. 2020.
- [53] Chen KT, Gong E, de Carvalho Macruz FB, Xu J, Boumis A, Khalighi M. Ultra-low-dose 18F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs. *Radiology*. 2019;290(3):649–56.
- [54] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B,

- Warde-Farley D, Ozair S. Generative adversarial nets. *Adv Neural Inf Processing Systems*. 2014:2672-80.
- [55] Wolterink JM, Dinkla AM, Savenije MH, Seevinck PR, van den Berg CA, Išgum I. Deep MR to CT synthesis using unpaired data. In *International Workshop on Simulation and Synthesis in Medical Imaging*. 2017 Sep:14-23.
- [56] Yi X, Babyn P. Sharpness-aware low-dose CT denoising using conditional generative adversarial network. *J Digit Imaging*. 2018;31(5):655-69.
- [57] Sánchez I, Vilaplana V. Brain MRI super-resolution using 3D generative adversarial networks. *arXiv preprint arXiv:1812.11440*; 2018.
- [58] Ran M, Hu J, Chen Y, Chen H, Sun H, Zhou J. Denoising of 3D magnetic resonance images using a residual encoder-decoder Wasserstein generative adversarial network. *Medl Image Anal*. 2019; 55:165-80.
- [59] Nie D, Trullo R, Lian J, Petitjean C, Ruan S, Wang Q. Medical image synthesis with context-aware generative adversarial networks. In *International Conference on Medical Image Computing and Computer-assisted Intervention*. 2017 Sep:417-25.

구분	성명	소속	직위
주저자	홍준용	동서대학교 대학원 융합방사선학과	대학원생
공동저자	박상현	동서대학교 대학원 융합방사선학과	대학원생
교신저자	정영진	동서대학교 방사선학과	부교수