

Comparative Evaluation of Chest Image Pneumonia based on Learning Rate Application

Ji-Yul Kim^{1,2}, Soo-Young Ye^{2,*}

¹Daewoo general hospital

²Department of Radiology Catholic University of the Pusan

Received: September 06, 2022. Revised: October 27, 2022. Accepted: October 31, 2022.

ABSTRACT

This study tried to suggest the most efficient learning rate for accurate and efficient automatic diagnosis of medical images for chest X-ray pneumonia images using deep learning. After setting the learning rates to 0.1, 0.01, 0.001, and 0.0001 in the Inception V3 deep learning model, respectively, deep learning modeling was performed three times. And the average accuracy and loss function value of verification modeling, and the metric of test modeling were set as performance evaluation indicators, and the performance was compared and evaluated with the average value of three times of the results obtained as a result of performing deep learning modeling. As a result of performance evaluation for deep learning verification modeling performance evaluation and test modeling metric, modeling with a learning rate of 0.001 showed the highest accuracy and excellent performance. For this reason, in this paper, it is recommended to apply a learning rate of 0.001 when classifying the presence or absence of pneumonia on chest X-ray images using a deep learning model. In addition, it was judged that when deep learning modeling through the application of the learning rate presented in this paper could play an auxiliary role in the classification of the presence or absence of pneumonia on chest X-ray images. In the future, if the study of classification for diagnosis and classification of pneumonia using deep learning continues, the contents of this thesis research can be used as basic data, and furthermore, it is expected that it will be helpful in selecting an efficient learning rate in classifying medical images using artificial intelligence.

Keywords: Deep learning, Learning-rate, Automatic Diagnosis of Pneumonia, Chest X-ray imaging, Inception V3

I. INTRODUCTION

2019년도에 발생한 집단 폐렴 병원체인 코로나바이러스-19의 전 세계적인 확산으로 인하여 확진자와 사망자가 기하급수적으로 증가하는 상황이다^[1]. 이로 인하여 의료진 및 의료시설의 부족에 대한 문제 해결과 코로나바이러스-19 감염의 확산 방지를 위한 신속하고 정확한 진단 방법의 필요성이 대두되고 있다^[2]. 2020년도에 보고된 임상미생물학 감염 학회의 보고에 의하면 폐렴은 호흡기 질환 중 사망자 수 1위에 해당하는 질병이다^[3]. 코로나바이러스-19에 감염된 환자들은 흉부 X선 영상을 이용

하여 폐렴 소견을 진단할 수 있으며^[2] X선 영상은 코로나바이러스-19 발생 이후 중요한 진단 검사 중 하나로 인식되고 있다^[4]. 하지만 실제 흉부 X선 영상 판독 시 판독자 간의 주관적인 해석의 가능성과 상당한 시간이 소요될 수 있다^[5].

현재의 의료기술은 나날이 갈수록 발전되어가고 있으나 질병의 진단은 여전히 환자들의 검사 결과를 바탕으로 의사들이 판단하는 전통적인 방법에 의존하고 있다^[6]. 하지만 최근에 인공지능 기술이 급격하게 발전함으로써 의료영상 분야에서는 인공지능을 이용하여 지각 정보를 해석하는 딥러닝 (Deep Learning) 기술이 활발하게 연구되고 있다^[7].

* Corresponding Author: Soo-Young Ye E-mail: syye@cup.ac.kr Tel: +82-51-510-0586

딥러닝은 인간의 신경계를 모방한 인공신경망(Artificial Neural Network)을 기반으로 설계된 개념이다^[8]. 여러 가지 형태의 인공신경망 중에서 컨볼루션 인공신경망(Convolutional Neural Network)은 필터링 기법을 적용하여 이미지를 효과적으로 처리할 수 있으며^[9] 의료영상 분야에서 기본적인 인공 신경망으로 사용되고 있다^[10]. 또한 딥러닝 기술은 의료 비용 상승의 저하, 의료 불균등 해소, 인력 부족 문제의 해결, 코로나로 인한 원격진료 등의 다양한 형태로도 발전하고 있다^[11].

모델링 실행은 딥러닝을 이용한 의료영상 분야에서 정확하고 효율적인 의료영상의 자동진단을 위해서 학습률(Learning rate)을 적용하였다^[12].

학습률이란 출력값의 오차를 줄이기 위한 경사하강법의 적용과정에서 학습 속도를 조절하기 위한 하이퍼 파라미터이다^[12]. 학습률이 너무 낮거나 높게 설정한다면 최적의 가중치를 찾는 데 오랜 시간이 걸리거나 발산한 값을 찾지 못하는 경우가 발생할 수 있다^[13]. 이러한 학습률은 딥러닝 모델링 시에 대체적으로 임의로 학습률을 설정하여 모델링을 실행한다.

본 논문은 정확하고 효율적인 흉부 X선 영상의 폐렴 유무 진단 분류를 위해 작은 수의 파라미터를 사용하여 연산량의 감소 효과를 기대할 수 있는 Inception V3 딥러닝 모델을 이용하였다. 각각의 학습률 적용에 따른 흉부 X선 영상의 폐렴 유무 진단 분류에 대한 성능을 비교 평가하고자 한다. 그리고 본 논문의 결과를 통하여 흉부 X선 영상의 폐렴 유무 진단 분류에 대한 최적의 학습률을 파악하고 향후 유사 연구 및 관련 산업에 최적의 학습률을 제시하고자 한다.

II. MATERIAL AND METHODS

1. 실험 재료

1.1. 딥러닝 모델링 환경

본 연구에서 딥러닝 모델링을 구현하기 위한 환경설정은 Table 1과 같다. 프로그래밍 언어로는 파이썬, 프레임워크는 케라스, 데이터세트는 Kaggle

에서 제공하는 흉부 X선 영상 데이터 세트를 이용하였다. 그리고 가상환경으로 아나콘다를 적용하였으며 주피터 노트북을 활용하여 파이썬 환경을 구축하였다.

Table 1. Environment of Deep learning

Programming language	Python (Version 3.10.0 Python Software Foundation, Delaware, USA)
Frame work	Keras (Version 2.4)
Dataset	Kaggle chest image dataset (license: CC BY 4.0)
Virtual environment	Anaconda (Version 4.10.1, Continuum Analytics, Texas, USA)

1.2. 합성곱 신경망

합성곱 신경망은 입력한 값으로부터 이미지의 고유한 특징 맵을 만들고 신경망에 연결한다^[14]. 그리고 각각의 해당되는 레이블에 이미지를 분류한다^[14]. 합성곱 신경망은 입력과 출력 부분에서 뉴런들이 학습해야 하는 가중치의 수가 작고 상대적으로 훈련 시간 대비 높은 성능을 나타낸다^[15]. 이러한 이유로 이미지 데이터 분류 기법에 주로 적용되고 있다^[15].

1.3. Inception V3 딥러닝 모델

Inception V3 딥러닝 모델 구조는 Fig. 1과 같다. 기존의 Inception 모델에서 수행하는 7x7 컨볼루션 연산을 1x1 컨볼루션 연산과 3x3 컨볼루션 연산으로 변환하여 수행하는 것이 특징적이다^[16]. 이러한 특징으로 인하여 Inception V3 딥러닝 모델은 Layer 증가에 연산량 및 소요 시간을 감소시킨다. 그리고 연산의 복잡도를 줄이고 파라미터 수를 적게 함으로써 성능 저하를 방지한다는 장점이 있다^[17].

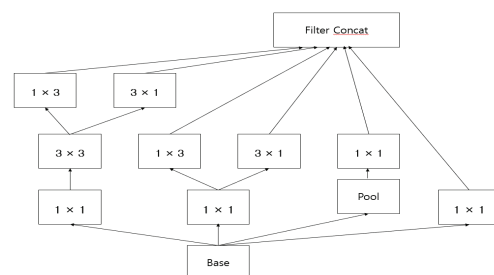


Fig. 1. Structure of Inception V3.

1.4. 학습률

학습률은 딥러닝 모델링을 최적화하기 위한 하이퍼 파라미터이다. 딥러닝 모델링 시 최적의 학습률 적용을 통하여 딥러닝 모델의 정확성과 효율성을 향상시킬 수 있다^[12].

2. 실험 방법

흉부 X선 영상을 이용한 폐렴 유무 자동 진단을 위해 Inception V3 딥러닝 모델을 선정하였으며, 동일 딥러닝 모델에 학습률을 0.1, 0.01, 0.001, 0.0001로 각각 설정한 다음 모델링을 3회 실시하였다.

2.1. 탐색적 데이터 분석

탐색적 데이터 분석은 데이터를 시각화하는 과정을 의미한다^[18]. 본 연구는 우선적으로 Kaggle에서 제공하는 흉부 X선 영상 데이터에 대한 탐색적 데이터 분석을 시행하였다. 흉부 X선 영상의 폐렴 유무 진단 분류에 대한 딥러닝 모델링을 구현하기 위하여 정상 흉부 X선은 라벨링 0, 폐렴 흉부 X선 영상은 라벨링 1로 라벨링을 하였다. 탐색적 데이터 분석 결과는 Table 2와 같다.

Table 2. Classification of dataset

	Labeling : 0 (Normal)	Labeling : 1 Pneumonia	Total
Training dataset	790	2140	2,930
Validation dataset	593	1833	2,426
Test dataset	200	300	500
Total	1,583	4,273	5,856

2.2. 학습 모델링

흉부 X선 영상의 폐렴 유무 분류에 대한 신속성과 정확성을 평가하기 위해서 Inception V3 딥러닝 모델에 적용한 파라미터는 Table 3과 같다.

Inception V3 딥러닝 모델링의 입력 이미지 사이즈는 224×224, 모델링의 시도 횟수는 60회, 출력층의 활성화 함수는 회귀분석 목적의 Sigmoid 함수, 배치 사이즈는 8로 설정하였다. 그리고 학습률은 0.1, 0.01, 0.001, 0.0001로 각각 적용하여 3회 딥러닝 모델링을 실행한 후 출력된 결과값의 평균을 산출하였다.

Table 3. Construction of Deep learning modeling

Input size	224 × 224			
Epoch	60			
Activation function	Sigmoid			
Batch size	8			
Optimizer	Adam			
Learning rate	0.1,	0.01,	0.001,	0.0001

3. 성능 평가

Inception V3 딥러닝 모델링의 흉부 X선 영상의 폐렴 유무 분류에 대한 학습률에 따른 결과를 비교 평가를 하기 위하여 검증 모델링의 정확도, 손실 함수의 결과값을 3회 평가하였다. 그리고 Test 모델링에 대한 Test Metric 정확도, 정밀도, 재현율, F1-score, AUC Score의 3회 실시한 Test 모델링의 평균값을 산출하여 평가하였다.

3.1. 정확도 평가

정확도란 예측값과 실제 값의 비교 시 정확한 정도를 나타내며, 가장 직관적인 모델의 성능평가 지표이다^[19].

Eq. (1)은 정확도를 구하는 수식이다^[20].

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

TP : True Positive
 TN : True Negative
 FP : False Positive
 FN : False Negative

3.2. 정밀도(precision) 평가

정밀도란 모델이 Positive라고 예측한 값 중에서 실제 Positive의 비율을 나타낸 것이다^[19].

Eq. (2)는 정밀도를 구하는 수식이다^[20].

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

3.3. 재현율(Recall) 평가

재현율이란 예측 값과 실제 값과의 비교 시 True인 것 중에서 모델이 Positive라고 예측한 것의 비율을 나타낸 것이다^[19].

Eq. (3)은 재현율을 구하는 수식이다^[20].

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

3.4. F1-score 평가

F1-score란 정밀도와 재현율의 조화평균이다^[19].

Eq. (4)는 F1-score를 구하는 수식이다^[20].

$$F1 - score = 2 \times \frac{1}{\frac{1}{Precision} + \frac{1}{Recall}} \quad (4)$$

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

Precision : 정밀도
Recall : 재현율

3.5. AUC score 평가

AUC score란 ROC(receiver operating characteristic) curve를 수치화한 것으로 1에 가까운 값일수록 좋은 모델이다^[18].

III. RESULT

본 연구는 흉부 X선 영상의 폐렴 유무 분류에 대한 최적의 학습률을 평가하였다. 이를 위해 모델링은 Inception V3 모델을 이용하여 각각의 학습률 별로 3회 실행하였다. 모델링 결과값은 3회 평균값을 비교 평가하였다.

1. 검증 모델링 평가

Inception V3 딥러닝 모델을 이용한 각각의 학습률 별 검증 모델링에 대한 정확도 결과 및 손실함수 결과를 비교 평가하였다.

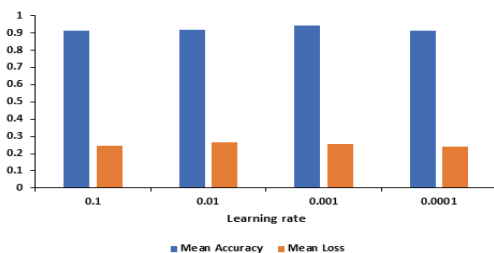


Fig. 2. Validation modeling mean result.

1.1. 정확도 평가

Inception V3 딥러닝 모델을 이용한 각각의 학습률 별 검증 모델링에 대한 정확도 결과값 및 3회 평균값은 Table 4와 같다.

Table 4. Validation modeling accuracy result

Learning rate	Validation Accuracy				
	1	2	3	Mean	Standard deviation
0.1	0.9198	0.9094	0.9156	0.9149	0.005
0.01	0.9206	0.9123	0.9181	0.9170	0.004
0.001	0.9919	0.9181	0.9214	0.9438	0.042
0.0001	0.9115	0.9256	0.9036	0.9134	0.011

검증 모델링에 대한 평균 정확도 평가는 학습률 0.1의 정확도 평균값은 0.9149, 학습률 0.01의 정확도 평균값은 0.9170, 학습률 0.001의 정확도 평균값은 0.9438, 학습률 0.0001의 정확도 평균값은 0.9135를 나타내었다. 이러한 결과로 학습률을 0.001로 적용한 딥러닝 모델링이 흉부 X선 영상의 폐렴 진단 유무를 분류하는 데 있어 가장 우수한 모델링이라고 평가하였다.

1.2. 손실함수 값 평가

Inception V3 딥러닝 모델을 이용한 각각의 학습률 별 검증 모델링에 대한 손실함수 결과값 및 3회 평균값은 Table 5와 같다.

Table 5. Validation modeling loss result

Learning rate	Validation Loss				
	1	2	3	Mean	Standard deviation
0.1	0.2374	0.2494	0.2553	0.2473	0.009
0.01	0.2532	0.2716	0.2736	0.2661	0.011
0.001	0.2587	0.2736	0.2367	0.2563	0.019
0.0001	0.2707	0.2141	0.2422	0.2423	0.028

검증 모델링에 대한 평균 손실함수 평가는 학습률 0.1의 손실 함수 평균값은 0.2473, 학습률 0.01의 손실 함수 평균값은 0.2661, 학습률 0.001의 손실 함수 평균값은 0.2563, 학습률 0.0001의 손실 함수 평균값은 0.2423을 나타내었다. 이러한 결과로 학습률을 0.0001로 적용한 딥러닝 모델링이 흉부 X선

영상의 폐렴 진단 유무를 분류하는 데 있어 인공지능 경망의 학습 상태가 가장 우수한 모델링이라고 평가하였다.

2. 테스트 모델링 Metric 평가

Inception V3 딥러닝 모델을 이용한 각각의 학습률 별 테스트 모델링에 대한 Metric 평가 결과의 3회 평균값은 Table 6과 같다.

Table 6. Test Metric result

Learning rate	Validation				
	Accuracy (%)	Precision (%)	Recall (%)	F1-score	AUC-score
0.1	94.47	100.0	90.77	95.15	0.9981
0.01	95.86	98.95	94.10	96.45	0.9959
0.001	94.85	99.88	91.55	95.53	0.9976
0.0001	93.39	99.49	89.44	94.19	0.9946

2.1. 정확도(Accuracy) 평가

학습률 0.1의 정확도는 94.47%, 학습률 0.01의 정확도는 95.86%, 학습률 0.001의 정확도는 94.85%, 학습률 0.0001의 정확도는 93.39%를 나타내었다. 이러한 결과로 학습률 0.01을 적용한 딥러닝 모델링이 흉부 X선 영상의 폐렴 진단 유무를 분류함에 있어 정확도가 가장 높은 모델링으로 평가하였다. 그리고 모든 학습률의 정확도 결과가 93%를 상회하는 결과를 나타내어 정확도 평가에서 우수하다고 판단하였다.

2.2. 정밀도(precision) 평가

학습률 0.1의 정밀도는 100%, 학습률 0.01의 정밀도는 98.95%, 학습률 0.001의 정밀도는 99.88%, 학습률 0.0001의 정밀도는 99.49%를 나타내었다. 이러한 결과로 학습률 0.1을 적용한 딥러닝 모델링이 흉부 X선 영상의 폐렴 진단 유무를 분류함에 있어 정밀도가 가장 높은 모델링으로 평가하였으며, 모든 학습률의 정밀도 결과가 98%를 상회하는 결과를 나타내어 정밀도 평가에서 매우 우수하다고 판단하였다.

2.3. 재현율(Recall) 평가

학습률 0.1의 재현율은 90.77%, 학습률 0.01의 재현율은 94.10%, 학습률 0.001의 재현율은 91.55%,

학습률 0.0001의 재현율은 89.44%를 나타내었다 (Table. 4). 이러한 결과로 학습률 0.01을 적용한 딥러닝 모델링이 흉부 X선 영상의 폐렴 진단 유무를 분류함에 있어 재현율이 가장 높은 모델링으로 평가하였다. 그리고 모든 학습률의 재현율 결과가 90%에 근접하거나 상회하는 결과를 나타내어 재현율 평가에서 우수하다고 판단하였다.

2.4. F1-score 평가

학습률 0.1의 F1-score는 95.15%, 학습률 0.01의 F1-score는 96.45%, 학습률 0.001의 F1-score는 95.53%, 학습률 0.0001의 F1-score는 94.19%를 나타내었다(Table. 4). 이러한 결과로 학습률 0.01을 적용한 딥러닝 모델링이 흉부 X선 영상의 폐렴 진단 유무를 분류함에 있어 F1-score가 가장 높은 모델링으로 평가하였다. 그리고 모든 학습률의 F1-score 결과가 95%에 근접하거나 상회하는 결과를 나타내어 F1-score 평가에서 매우 우수하다고 판단하였다.

2.5. AUC score 평가

학습률 0.1의 AUC score는 0.9981, 학습률 0.01의 AUC score는 0.9959, 학습률 0.001의 AUC score는 0.9976, 학습률 0.0001의 AUC score는 0.9946를 나타내었다. 이러한 결과로 모든 학습률을 적용한 딥러닝 모델링은 흉부 X선 영상의 폐렴 진단 유무를 분류함에 있어 AUC score가 가장 높은 모델링으로 평가하였다. 그리고 모든 학습률의 AUC score 결과가 99%를 상회하는 결과를 나타내어 AUC score 평가에서 매우 우수하다고 판단하였다.

IV. DISCUSSION

딥러닝에서의 최적화란 인공지능경망 모델이 학습과 그 결과에 따른 손실함수의 값을 최소화하는 방향으로 하이퍼 파라미터의 값을 찾는다는 것을 의미한다. 딥러닝 모델링의 정확도와 효율성을 향상시키기 위해서는 딥러닝 모델링에 의료영상을 학습시키는 과정에서 학습률 설정이 매우 중요하며, 가장 효과적인 학습률 값을 찾는 것을 목표로 한다.

2020년도에 보고된 이종근 등의 연구에서는 흉부 X선 영상에 딥러닝 알고리즘을 적용하여 병변의 유무를 판단하고자 하였다. 학습률 0.0001을 적

용하여 평가지표로 AUC를 사용하였으며 그 결과 제안한 알고리즘의 AUC score가 정상 흉부영상에서는 0.97, 비정상 흉부영상에서는 0.931의 결과를 나타내었다^[21].

본 연구는 흉부 X선 영상에 대한 이용한 폐렴 유무 자동진단을 위해 최적의 학습률을 평가하고자 Inception V3 모델을 이용하여 모델링을 실행하였다. 학습률을 0.1, 0.01, 0.001, 0.0001로 각각 설정하여 모델링 한 결과 학습률 0.1의 AUC score는 0.9981, 학습률 0.01의 AUC score는 0.9959, 학습률 0.001의 AUC score는 0.9976, 학습률 0.0001의 AUC score는 0.9946를 나타내었다. 이러한 결과로 본 연구에서 적용한 딥러닝 모델링의 AUC score 평가가 더 우수하다고 판단하였다.

2021년도에 보고된 정우연 등의 연구에서는 Inception V3 딥러닝 모델을 이용하여 흉부 X선 영상의 심장 비대증을 분류하고자 했으며 심장비대증의 특징 추출의 경우 0.0001의 학습률을 적용하였다^[22]. 본 연구는 흉부 X선 영상을 이용한 폐렴 유무 자동진단을 위해 Inception V3 딥러닝 모델 학습률을 0.1, 0.01, 0.001, 0.0001로 각각 설정하여 모델링 한 결과 학습률 0.001을 적용한 딥러닝 모델링이 성능평가나 지표평가에서 매우 우수한 결과를 나타내었다. 본 연구를 통하여 향후 흉부 X선 영상을 이용한 딥러닝 모델링의 심장 비대증 분류 시 최적의 학습률을 제시하기 위한 추가적인 연구의 필요성을 인식하였다.

V. CONCLUSION

본 논문은 Inception V3 딥러닝 모델링을 이용하여 흉부 X선 영상을 사용한 폐렴의 진단 유무 분류에 다양한 학습률 별 딥러닝 모델링을 실행하여 가장 효과적인 학습률 값을 확인하고자 하였다. 딥러닝 검증 모델링 성능평가 및 Test 모델링 Metric에 대한 성능평가의 결과, 학습률 0.001을 적용한 모델링이 가장 높은 정확도와 우수한 성능을 나타내었다. 이러한 이유로 본 논문은 딥러닝 모델을 이용한 흉부 X선 영상에 대한 폐렴 유무 분류 시 학습률을 0.001로 적용할 것을 권고한다. 그리고 본 논문에서 제시하는 학습률의 적용을 통한 딥러닝 모

델링 시 흉부 X선 영상에 대한 폐렴 유무 분류에 대한 인력의 보조적인 역할을 수행할 수 있을 거라고 판단하였다.

향후 딥러닝을 이용한 폐렴 유무 진단 분류 연구가 계속해서 진행될 시, 본 논문의 논문 연구 내용은 기초자료로 활용될 수 있다고 여겨지며 나아가 인공지능을 활용한 의료영상 분류에 있어 효율적인 학습률 선택에 도움이 될 것으로 기대된다. 그리고 본 연구는 폐렴 유무 진단의 분류뿐만 아니라 폐렴의 검출 및 분할 등의 추가적인 연구의 시작점이 될 것이라고 기대한다.

Acknowledgement

본 연구는 2022년 부산가톨릭대학교 교내학술연구과제로 수행되었다.

Reference

- [1] K. H. Ann, S. Y. OUM, "A COVID-19 Chest X-ray Reading Technique based on Deep Learning", The International Promotion Agency of Culture Technology, Vol. 6, No. 4, pp. 789-795, 2020. <http://dx.doi.org/10.17703/JCCT.2020.6.4.789>
- [2] S. E. Park, "Epidemiology, virology, and clinical features of severe acute respiratory syndrome -coronavirus-2 (SARS-CoV-2; Coronavirus Disease-19)", Clinical and Experimental Pediatrics, Vol. 63, No. 4, pp. 119-124, 2020. <https://doi.org/10.3345/cep.2020.00493>
- [3] N. Petrosillo, G. Viceconte, O. Ergonul, G. Ippolito, E. Petersen, "COVID-19 SARS and MERS: are they closely related?", Clinical Microbiology and Infection, Vol. 26, No. 6, pp. 729-734, 2020. <http://dx.doi.org/10.1016/j.cmi.2020.03.026>
- [4] Y. J. Kim, K. G. Kim, "Development of an Optimized Deep Learning Model for Medical Imaging", Journal of The Korean Society of Radiology, Vol. 81, No. 6, pp. 1274-1289, 2020. <https://doi.org/10.3348/jksr.2020.0171>
- [5] L. Yao, E. Poblentz, D. Dagunts, B. Covington, D. Bernard, K. Lyman, "Learning to diagnose from scratch by exploiting dependencies among labels", 2017. <https://doi.org/10.48550/arXiv.1710.10501>

- [6] J. Kim, "An Analysis of the effect of Artificial Intelligence on Human Society", *The Journal of the Convergence on Culture Technology*, Vol. 5, No. 2, pp. 177-182, 2019. <https://doi.org/10.17703/JCCT.2019.5.2.177>
- [7] T. H. Lim, K. O. Lim, S. W. Chung, S. C. Han, "Disease diagnosis research using deep learning based on military medical data", *Journal of Digital Contents Society*, Vol. 22, No. 9, pp. 1359-1367, 2021. <http://doi.org/10.9728/dcs.2021.22.9.1359>
- [8] K. W. Yang, J. S. Kang, H. N. Lee, "Digital Walkie-Talkie Identification scheme based on Convolutional Neural Network", *The Journal of Korean Institute of Communications and Information Sciences*, pp. 1210-1211, 2019.
- [9] G. Litjens, et al, "A survey on deep learning in medical image analysis", *Medical Image Analysis*, Vol. 42, pp. 60-88, 2017. <http://dx.doi.org/10.1016/j.media.2017.07.005>
- [10] A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet classification with deep convolutional neural networks", *Communications of the ACM*, Vol. 60, No. 6, pp. 84-90, 2017. <http://dx.doi.org/10.1145/3065386>
- [11] M. Rastgarpour, J. Shanbehzadeh, "Application of AI Techniques in Medical Image Segmentation and Novel Categorization of Available Methods and Tools", *Lecture Notes in Engineering and Computer Science*, Vol. 2188, No. 1, pp. 519-523, 2011. <https://doi.org/10.2174/1573405617666210713113439>
- [12] J. H. Lee, B. S. Kang, "Improving Performance of Machine Learning-Based Algorithms with Adaptive Learning Rate", *The Journal of Korean Institute of Information Technology*, Vol. 18, No. 10, pp. 9-14, 2020. <http://doi.org/10.14801/jkiit.2020.18.10.9>
- [13] M. J. Kang, "Comparison of Gradient Descent for Deep Learning", *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 21, No. 2 pp. 189-194, 2020. <https://doi.org/10.5762/KAIS.2020.21.2.189>
- [14] H. J. Song, E. B. Lee, H. J. Jo, S. Y. Park, S. Y. Kim, H. J. Kim, J. W. Hong, "Evaluation of Classification and Accuracy in Chest X-ray Images using Deep Learning with Convolution Neural Network", *Journal of the Korean Society of Radiology*, Vol. 14, No. 1, pp. 39-43, 2020. <http://doi.org/10.7742/jksr.2020.14.1.39>
- [15] S. J. Kim, J. C. Yu, "COVID-19 Chest X-ray reading Technique based on Deep Learning", *Proceedings of the Korean Society of Computer Information Conference*, Vol. 29, No. 1, pp. 31-32, 2021.
- [16] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, "Rethinking the Inception Architecture for Computer Vision", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818-2826, 2016.
- [17] N. Dong, L. Zhao, C. H. Wu, J. F. Chang, "Inception v3 based cervical cell classification combined with artificially extracted features", *Applied Soft Computing*, Vol. 93, pp. 106311, 2020. <http://dx.doi.org/10.1016/J.ASOC.2020.106311>
- [18] <https://eda-ai-lab.tistory.com/13>
- [19] K. D. Song, M. C. Kim, S. H. Do, "The Latest Trends in the Use of Deep Learning in Radiology Illustrated Through the Stages of Deep Learning Algorithm Development", *Korean Journal of Radiology*, Vol. 80, No. 2, pp. 202-212, 2019. <http://dx.doi.org/10.3348/jksr.2019.80.2.202>
- [20] H. Huhtanen, M. Nyman, T. Mohsen, A. Virkki, A. Karlsson, J. Hirvonen, "Automated detection of pulmonary embolism from CT-angiograms using deep learning", *BMC Medical Imaging*, Vol. 22, 2022. <http://dx.doi.org/10.1186/s12880-022-00763-z>
- [21] J. K. Lee, "Detection and Classification of Pneumothorax and Pneumonia through Deep Learning in Chest X-ray", *Graduate of National university of Chung-buk, Master's thesis*, 2020.
- [22] W. Y. Jeong, J. H. Kim, J. E. Park, M. J. Kim, J. M. Lee, "Evaluation of Classification Performance of Inception V3 Algorithm for Chest X-ray Images of Patients with Cardiomegaly", *Journal of the Korean Society of Radiology*, Vol. 15, No. 4, pp. 455-461, 2021. <https://doi.org/10.7742/jksr.2021.15.4.455>

학습률 적용에 따른 흉부영상 폐렴 유무 분류 비교평가

김지율^{1,2}, 예수영^{2,*}

¹대우병원 영상의학과

²부산가톨릭대학교 방사선학과

요 약

본 연구는 딥러닝을 이용한 흉부 X선 폐렴 영상에 대하여 정확하고 효율적인 의료영상의 자동진단을 위해서 가장 효율적인 학습률을 제시하고자 하였다.

Inception V3 딥러닝 모델에 학습률을 0.1, 0.01, 0.001, 0.0001로 각각 설정한 후 3회 딥러닝 모델링을 수행하였다. 그리고 검증 모델링의 평균 정확도 및 손실 함수 값, Test 모델링의 Metric을 성능평가 지표로 설정하여 딥러닝 모델링의 수행 결과로 획득한 결과값의 3회 평균값으로 성능을 비교 평가하였다.

딥러닝 검증 모델링 성능평가 및 Test 모델링 Metric에 대한 성능평가의 결과, 학습률 0.001을 적용한 모델링이 가장 높은 정확도와 우수한 성능을 나타내었다. 이러한 이유로 본 논문에서는 딥러닝 모델을 이용한 흉부 X선 영상에 대한 폐렴 유무 분류 시 학습률을 0.001로 적용할 것을 권고한다. 그리고 본 논문에서 제시하는 학습률의 적용을 통한 딥러닝 모델링 시 흉부 X선 영상에 대한 폐렴 유무 분류에 대한 인력의 보조적인 역할을 수행할 수 있을 거라고 판단하였다.

향후 딥러닝을 이용한 폐렴 유무 진단 분류 연구가 계속해서 진행될 시, 본 논문의 논문 연구 내용은 기초자료로 활용될 수 있다고 여겨지며 나아가 인공지능을 활용한 의료영상 분류에 있어 효율적인 학습률 선택에 도움이 될 것으로 기대된다.

중심단어: 딥러닝, 학습률, 폐렴, 폐렴 유무 자동진단, 흉부 X선 영상, Inception V3

연구자 정보 이력

	성명	소속	직위
(제1저자)	김지율	대우병원 영상의학과 부산가톨릭대학교 방사선학과	방사선사 강사
(교신저자)	예수영	부산가톨릭대학교 방사선학과	부교수