

CT영상에서의 AlexNet과 VggNet을 이용한 간암 병변 분류 연구

최보혜¹ · 김영재² · 최승준³ · 김광기²

¹가천대학교 보건과학대학 의용생체공학과, ²가천대학교 의과대학 의공학교실, ³가천대학교 길병원 영상의학과

Malignant and Benign Classification of Liver Tumor in CT according to Data pre-processing and Deep running model

Bo Hye Choi¹, Young Jae Kim², Seung Jun Choi³ and Kwang Gi Kim²

¹Bio-Medical Engineering, College of Health Science, Gachon University

²Department of Biomedical Engineering, College of Medicine, Gachon University

³Department of Radiology, Gachon University Gil Hospital

(Manuscript received 18 June 2018 ; revised 25 October 2018 ; accepted 30 October 2018)

Abstract: Liver cancer is one of the highest incidents in the world, and the mortality rate is the second most common disease after lung cancer. The purpose of this study is to evaluate the diagnostic ability of deep learning in the classification of malignant and benign tumors in CT images of patients with liver tumors. We also tried to identify the best data processing methods and deep learning models for classifying malignant and benign tumors in the liver. In this study, CT data were collected from 92 patients (benign liver tumors: 44, malignant liver tumors: 48) at the Gil Medical Center. The CT data of each patient were used for cross-sectional images of 3,024 liver tumors. In AlexNet and VggNet, the average of the overall accuracy at each image size was calculated: the average of the overall accuracy of the 200 × 200 image size is 69.58% (AlexNet), 69.4% (VggNet), 150 × 150 image size is 71.54%, 67%, 100 × 100 image size is 68.79%, 66.2%. In conclusion, the overall accuracy of each does not exceed 80%, so it does not have a high level of accuracy. In addition, the average accuracy in benign was 90.3% and the accuracy in malignant was 46.2%, which is a significant difference between benign and malignant. Also, the time it takes for AlexNet to learn is about 1.6 times faster than VggNet but statistically no different ($p > 0.05$). Since both models are less than 90% of the overall accuracy, more research and development are needed, such as learning the liver tumor data using a new model, or the process of pre-processing the data images in other methods. In the future, it will be useful to use specialists for image reading using deep learning.

Key words: Liver cancer, Deep learning, AlexNet, VggNet, Classification

I. 서 론

Corresponding Author : Kwang Gi Kim
Gachon University College of Medicine, 21, Namdong-daero
774 beon-gil, Namdong-gu, Incheon, 21565, Republic of Korea
TEL: +82-32-820-4036
E-mail: kimkg@gachon.ac.kr
Corresponding Author: Seung Joon Choi
Department of Radiology, Gil Medical Center, Gachon
University School of Medicine, Incheon, South Korea
Tel: +82-32-460-3060
E-mail: sjchoi1118@gmail.com
연구는 국가과학기술연구회 융합연구과제(1711061488)의 지원을
받아 수행된 연구임.

간암은 세계적으로 발병률이 높은 암 가운데 하나이다. 간암의 발생률은 서양보다 동남아시아에서 발생률이 더 높다. 동남 아시아와 사하라 남쪽 아프리카 지역에서는 발생률이 가장 높아 10만 명당 30명의 환자가 발생한다고 보고되었다[1]. 간에 질병이 발생하면 환자는 초기 발병 사실을 거의 느끼지 못하다가 증상이 뚜렷해졌을 때에는 이미 질병이 진행된 단계가 대부분이다. 국립 암센터의 통계에 따르면, 2015년도 국내의 간암 발생률은 6번째이지만, 사망률은 폐

암에 이어서 가장 높은 질병이다. 또한 2011년부터 2015년까지 전체 암환자의 5년 상대생존율은 70.7%인 반면, 간암은 7대암 중 2번째로 생존율이 낮은 33.6%이다[1-2]. 따라서, 간암 환자의 조기 진단은 매우 중요하다. 보통 암을 진단할 때에는 조직검사를 실시하지만 간암을 진단하기 위해서는 조직검사 없이 전산화 단층 촬영(Computer Tomography, CT)이나 자기공명영상(Magnetic Resonance Imaging, MRI)과 같은 영상검사와 혈액검사(종양표지자 검사)로 진단한다. 이것으로 진단이 되지 않을 경우에는 조직검사까지 실시하게 된다. 대개 여러 차례의 영상검사를 통해 진단을 받게 되는데 이 검사를 통해 즉각적인 진단을 할 수 있다는 장점이 있다. 하지만 판독의사에 따라 암에 합당한 소견이 조금씩 달라질 수 있고 폐 종양(Tumor)의 경우에는 장기 주변 공기의 대비로 인해 육안으로도 구분이 가능할 정도이다. 하지만 간 종양의 경우에는 육안으로 구분하기가 비교적 어렵다. 그렇기 때문에 간암을 정확히 진단하고 확진할 수 있는 객관적인 진단방법이 필요하다[3].

최근에 컴퓨터 보조 진단에 대한 연구가 컴퓨터 능력 및 인공지능의 능력에 따라서 활용성이 높아졌다. 인공지능에 대한 분류 방식은 선형회귀분석(Logistic Regression Analysis), 베이지안(Bayesian), 랜덤 포레스트(Random Forest), 서포트 벡터 머신 (Support Vector Machine, SVM) 등에 의한 분류 방식으로 여러 연구들이 수행되어왔다[4-5]. 이러한 방식들은 뇌종양의 종류에 따라 분류하기 위해, 유방암의 악성과 양성을 구분하기 위해서 여러가지 형태로 의료계에 적용되어 연구되고 있다[6-8].

현재 딥러닝(Deep Learning) 관련해서 그래픽스 처리장치 (Graphics Processing Unit, GPU)를 이용한 인공지능(Artificial Intelligence) 계산 방식이 도입되었고, 이전에 수학적 연산 방식에 비해서 정확하고 빠른 결과를 도출한다. 최근 연구 방식으로는 안저 영상에서의 당뇨망막병증(Diabetic Retinopathy) 등의 검출 기법이 연구되었고, 폐종양(Pulmonary Tumors), 폐결핵(Pulmonary Tuberculosis), 대장 용종(Colon Polyp), 유방종괴(Breast Mass) 등에서 종양들의 양성과 악성들을 구분하는데 있어서도 인공지능이 도입되었다. 또한 의료 영상에 대해서 종양의 영상 분할, 심장의 영상 분할과 같은 결과 등이 많이 보고되었다[9-10]. 폐 결절 검출을 위해서 연세대학교 의과대학 의공학교실에서는 합성곱 신경망(Convolutional Neural Network, CNN)의 성능을 개선시켜 폐 결절 검출을 수행했고, 데이터증가기법(Data Augmentation)과 L1, L2 정규화 기법을 같이 적용한 합성곱 신경망에서 각각 93.66%와 93.75%의 정확도를 보였다. 하지만 제안된 합성곱 신경망에 과적합(Overfitting) 완화 방법을 적용하지 않고 학습하였을 때

의 정확도(Accuracy)는 91.79%로 나타났다[11]. 이주희 외 3명의 연구팀에서는 X-ray영상에서 폐결핵을 자동 진단하기 위해 SegNet을 이용한 유용성을 평가하는 연구를 진행하였다. 폐결핵을 위한 컴퓨터보조진단(Computer Aided Diagnosis, CAD) 시스템인 Breuninger's CAD4TB, Jaeger's Automatic TBScreening과 딥러닝 모델 중 SegNet을 이용하였을 때, CAD 시스템들은 각각 84.0%, 86.9%의 정확도를 보였으나, SegNet은 90.4%로 나타났다. 이 연구에서 사용된 데이터셋(Dataset)은 총 552장으로 데이터의 양이 부족했으나, 딥러닝이 보조진단으로 활용될 수 있다는 가능성을 보였다[12].

종양은 세포가 비정상적으로 증식해서 만들어진 덩어리로서, 인체에 해로운 영향을 준다. 특히, 악성종양(Malignant Tumor)의 경우에는 조기에 적절한 치료를 받지 않을 경우 생명을 위협하기 때문에 의사의 진단을 통해 빠른 치료를 받는 것이 매우 중요하다. 간의 악성종양 경우, 현재 우리나라 간암의 생존률은 7대암 중 2번째로 낮고, 사망률이 높다. 그렇기 때문에, 악성종양과 양성종양(Benign Tumor)을 빠른 시간 내에 분류하여 의사가 진단을 내리는 데에 빠른 도움을 주고 환자가 치료받을 수 있도록 하는 것이 매우 중요하다. 인체의 여러 장기들의 의료영상에서 인공지능을 활용한 연구가 활발히 진행되고 있는 가운데, 폐 종양이나 뇌 종양같은 질병에는 여러 딥러닝 모델 방식으로 적용되고 있지만 아직 간암에 대해서는 딥러닝을 활용한 연구나 결과가 부진하다. 그래서 본 연구는 간암 영상에서 딥러닝을 활용한 보조 진단 능력을 확인하고 간의 악성 종양과 양성 종양을 분류해내는 데에 가능성을 보고자 진행됐다.

관련 연구들의 여러 딥러닝 모델들 중 SegNet으로 폐결핵 자동 검출하는 데에 있어 가능성을 보인 기존 딥러닝 모델이 있으나 SegNet은 주로 분할(Segmentation)을 위한 모델로, 자율주행분야에서 주변 환경을 인식하기 위해 자주 쓰이는 모델이다. 그래서 본 연구에서는 간암의 양성과 악성을 분류하는 데에 AlexNet과 VggNet 모델을 활용하고자 하였다. 본 연구에서는 AlexNet과 VggNet 모델을 활용하여 양성과 악성 종양을 분류하는데 가장 적합한 데이터 처리 방법과 딥러닝 모델의 성능을 검증하고자 하였다. 기존 딥러닝 연구들에서는 하드웨어상의 제약이 있어 원본데이터를 그대로 사용하기에 어려움이 있었기 때문에 본 연구에서는 환자의 CT 이미지 데이터를 이미지 크기에 따라 나누어 이미지 크기를 다르게 함에 따라 가장 적절한 이미지 크기를 찾아내고자 하였다. 또한, 배치사이즈(Batch Size)와 에폭(Epoch)의 설정 값에 따라 여러 방식으로 처리하여 딥러닝 분류 모델에 따른 결과를 분석했다.

II. 재 료

본 연구에서는 가천대학교 길병원에서 악성 및 양성 간 종양이 발견된 92명(양성종양 :44명, 악성종양: 48명) 환자의 CT 데이터를 수집하였다. 수집된 각 환자의 CT데이터에서 간 종양이 위치한 3,024장(양성종양: 1,362장, 악성종양: 1,662장)의 횡단면 영상을 실험에 사용하였다. 각각의 양성과 악성 종양의 이미지 중에서 시험데이터(Test Data Set)를 각각 300장, 확인데이터(Validation Data Set)를 각각 200장으로 정하고, 훈련데이터(Training Data Set)에서 양성 종양 이미지는 862장, 악성 종양 이미지는 1162장을 사용하였다.

영상 전처리하는 전과정은 MATLAB(R2017b, The MathWorks Inc., Natick, USA)을 활용하여 작업을 수행하였고, AlexNet과 VggNet16 모델을 이용하기 위해 우분투(Ubuntu) OS에서 진행되었다. 두 개의 학습모델링을 사용하기 위해 텐서플로우(Tensorflow)를 활용한 코드로 구성하여 학습을 시켜 결과를 분석하였다.

III. 연구방법

1.영상 전처리 과정(Image Data Preprocessing)

CT 횡단면 영상들은 딥러닝을 활용하여 학습 및 시험하기 전에 그림1과 같이 이미지 전처리 과정을 수행하였다. 영상 전처리 과정에 따라 딥러닝의 학습결과가 달라질 수 있기 때문에 여러 방식으로 영상 전처리 과정을 수행하였다(그림1).

원본이미지는 의료용 디지털 영상 및 통신(Digital Imaging and Communications in Medicine, DICOM) 표준 포맷을 가지며 그림1에서 보드시피 12bit로 구성되어 있다. 조 직에 따라서 CT 영상 디스플레이에 사용되는 윈도우 폭(Window Width, WW)과 윈도우 레벨(Window Level, WL)이 다르다. 기본적으로 공기에 해당하는 픽셀은 -1000의 하운스필드단위(Hounsfield Unit, HU)값을 갖고 물에 해당하는 픽셀은 0HU를 갖는다[13]. 본 연구에서는 일반적으로 간 CT 영상 디스플레이에 사용되는 윈도우 폭과 윈도우 레벨을 조절하여 8bit이미지로 변환하였다. 본 연구에서는 일반적으로 임상에서 사용되는 윈도우 폭, 윈도우 레벨 각각 150HU, 50HU의 값을 사용하였다[14].

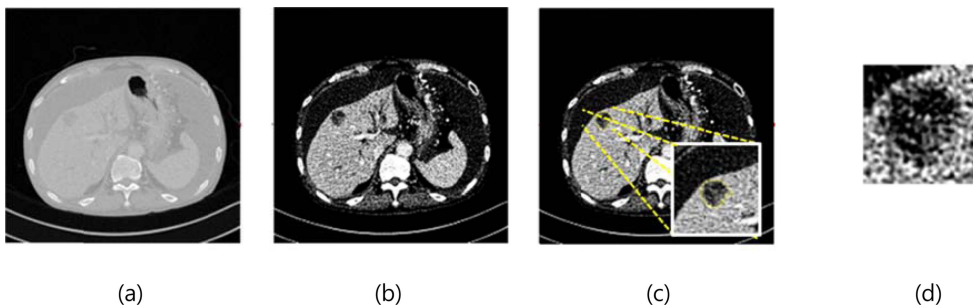


그림 1. 영상 전처리 과정의 예.

Fig. 1. Image data pre-processing process : (a) Original Image(12bit) (b) WW/WL(8bit) (c) ROI (d) Resize(150X150).

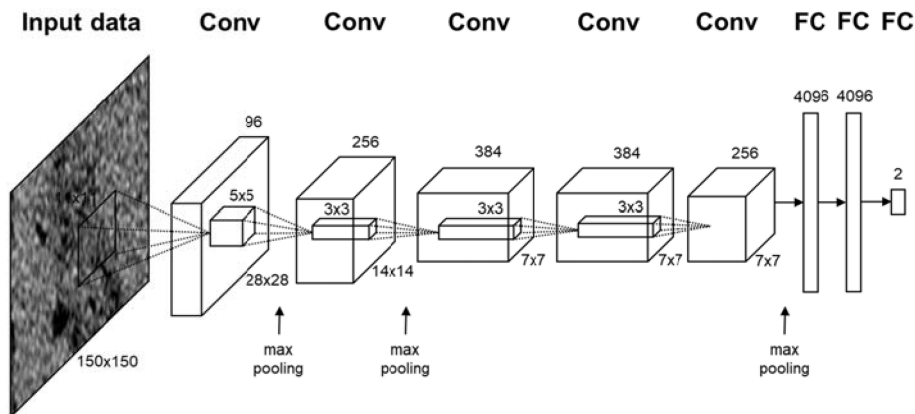


그림 2. AlexNet : 5개 컨볼루션 층, 3개의 완전 연결된 층.

Fig. 2. AlexNet : 5 convolutional layer, 3 fully-connected layer.

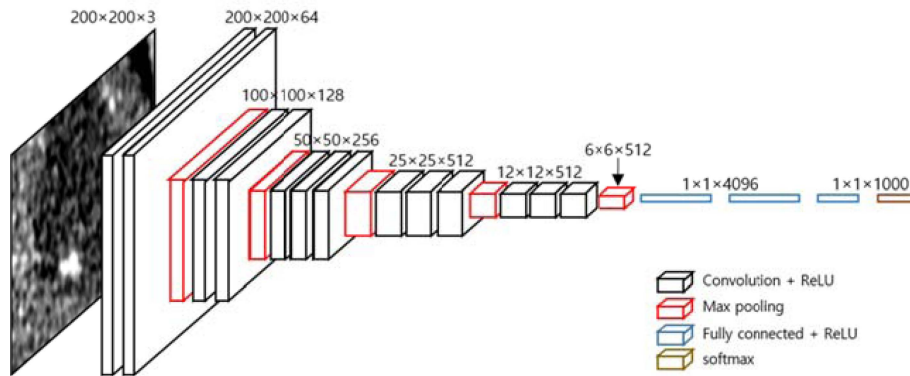


그림 3. Vgg Net 구조.
Fig. 3. Vgg Net Architecture.

윈도우 세팅(Window Setting)을 적용한 영상에서 전문 의의 도움을 받아 종양이 위치한 부분을 알아내고 Image J Program(National Institutes of Health, USA)를 사용하여 종양이 위치한 관심영역(Region Of Interest, ROI)을 표시하고 잘라내어 새로운 데이터로 만들었다. 환자에 따라서 관심영역을 추출한 종양의 크기가 각기 다르고, 이미지 크기에 따른 결과를 비교하기 위해 100x100, 150x150, 200x200의 크기로 변환(resampling)하였다.

232

2. 학습 모델링

본 연구에서는 이미지 분류를 위한 다양한 모델 중 AlexNet과 VggNet을 사용하였으며, 각 모델의 이미지 크기(Image size, img size)당 배치사이즈(batch size, bs)와 반복 횟수(epoch, ep)을 다르게 하여 각각의 결과를 획득하였다. 배치사이즈는 전체데이터 개수의 배치사이즈만큼 학습한 뒤 실제값과 비교하여 결과를 얻으면서 가중치가 갱신되는 것을 의미하며, 반복횟수는 학습을 위한 반복 횟수를 의미한다.

AlexNet은 2012년 ImageNet ILSVRC대회에서 2위와 큰 성능차를 보이며 우승을 한 것으로 알려진 CNN구조이다 [15]. AlexNet의 구조는 LeNet과 유사하지만 convolutional layer 다음에 pooling(sub-sampling) layer가 오는 기본 구조와 다르게 convolutional layer 바로 뒤에 convolutional layer가 온다(그림 2).

VggNet16은 대량의 이미지를 인식함에 있어서, 망의 깊이(depth)가 정확도에 어떤 영향을 주는지를 보여주었다 [16]. 망의 시작부터 끝까지 동일하게 3x3 convolution과 2x2 max pooling을 사용하는 단순한(homogeneous) 구조에서 depth가 16일때 최적의 결과를 보였다. Google에서 개발한 GoogLeNet에 비해 분류 성능은 약간 떨어지지만, 다중 전달 학습과제에서는 오히려 더 좋은 결과가 나온

구조이다. 단점은 메모리 수와 파라미터 수가 크다(그림 3).

본 연구에서 학습모델링에 따라 이미지 크기와 배치 사이즈, 에폭을 다르게 하여 학습시켰다. 이미지 크기는 전체 이미지들의 평균 크기를 기준으로 크게 3가지로 나누어 100x100, 150x150, 200x200로 결정하였고, 배치사이즈와 에폭은 각각 20, 50, 100(batch size)와 100, 800, 1000(epoch)으로 나누어 학습시켰다.

IV. 결 과

본 연구에서는 간암 CT 영상에서 간암의 악성과 양성 종양에 대하여 AlexNet과 VggNet모델을 활용하여 다양한 조건에서 학습을 시켰다. 결과는 크게 2가지로 나누어 확인하였다. 첫번째로 이미지 크기 및 배치 사이즈와 반복횟수의 차이에 따른 AlexNet과 VggNet에서 결과를 분석 및 비교하고, 두번째로는 두 모델을 활용하여 조건에 따라 학습하는데 걸리는 시간과 학습 결과 정확도를 확인 및 비교하였다.

1. 이미지 크기(Image size) 및 배치사이즈(batch size, bp)와 반복횟수(epoch, ep)에 차이에 따른 AlexNet과 VggNet 결과 비교 및 분석

이미지 크기 및 배치사이즈와 반복횟수의 차이에 따른 두 모델의 결과를 표1으로 나타냈다. 표1의 AlexNet에서 각 이미지 크기에서 배치사이즈와 반복횟수가 증가함에 따라 전체정확도(Overall Accuracy)가 작은 폭으로 계속 증가하고 있다. 악성의 정확도에서도 전체정확도와 같은 양상을 보이고 있으며 최대 60%(img size = 150, bs = 100, ep = 1000)의 정확도를 나타냈다. 하지만 양성의 정확도에서는 배치사이즈와 반복횟수가 증가함에 따라 증가하는 양상을 보이지 않고 랜덤적으로 경우에 따라 다르게 나타났다.

VggNet에서는 전체정확도의 결과가 AlexNet과는 달리,

표 1. 이미지 크기, 배치 크기, 반복 횟수의 차이에 따른 AlexNet과 VggNet 결과 비교

Table 1. Comparison of AlexNet and VggNet results due to difference in image size, batch size and epoch

Size (pixel)	Batch	Epoch	AlexNet			VGGNet		
			Overall Acc. (%)	Benign Acc. (%)	Malignant Acc. (%)	Overall Acc. (%)	Benign Acc. (%)	Malignant Acc. (%)
200	20	100	64.5	89.3	39.7	69.5	85.0	54.0
×	50	1000	68.3	94.0	42.7	67.7	90.0	45.3
200	100	1000	73.5	87.7	59.3	71.0	91.0	51.0
150	20	100	66.7	90.0	43.3	68.7	89.7	47.7
×	50	1000	70.3	91.0	49.7	66.0	93.7	38.3
150	100	1000	75.2	90.3	60.0	66.2	92.0	40.3
100	20	100	65.2	91.7	38.7	69.5	91.0	48.0
×	50	1000	68.3	89.7	47.0	63.3	89.7	37.0
100	100	1000	69.2	88.3	50.0	65.8	92.0	39.7

표 2. AlexNet과 VggNet의 학습 시간 비교 결과

Table 2. AlexNet and VggNet results through analysis of learning time

Size (pixel)	Batch	Epoch	AlexNet		VGGNet	
			Overall Acc. (%)	Time (seconds)	Overall Acc. (%)	Time (seconds)
200	20	100	64.5	1866.8	69.5	3597.652
×	50	1000	68.33	9040.4	67.7	10142.329
200	100	1000	73.5	5101.7	71.0	7881.351
150	20	100	66.7	1807.3	68.667	2710.1
×	50	1000	70.33	8679.8	66.0	8687.5
150	100	1000	75.17	4785.9	66.2	4886.695
100	20	100	65.167	1738.5	69.5	2150.4
×	50	1000	68.33	7582.9	63.333	11385.0
100	100	1000	69.167	3690.0	65.833	7280.1

각 이미지 크기에서 배치사이즈와 반복횟수가 증가함에 따라 전체정확도가 증가함을 보이지 않고 무작위적인 결과를 나타냈다. 200×200크기의 이미지에서는 배치사이즈가 100이고 반복횟수가 1000일 때 71%로 가장 높은 전체정확도를 나타냈고 그 외의 이미지 크기에서는 배치사이즈가 20이고 반복횟수가 100일때 전체정확도가 가장 높게 나타났다.

표1의 결과를 가지고 AlexNet에서 각 이미지 크기에서 전체정확도의 평균을 구하면, 200×200의 전체정확도의 평균은 69.58%, 150×150의 전체정확도의 평균은 71.54%, 100x100의 전체정확도의 평균은 68.79%를 나타낸다. 결론

적으로 이미지 크기 사이의 정확도는 거의 비슷한 수준의 값을 나타내고 있다. VggNet에서도 동일하게 전체정확도의 평균을 구하면, 각각 69.4% (img size = 200), 66.9% (img size = 150), 66.2% (img size = 100)로 AlexNet과 마찬가지로 거의 비슷한 수준의 값을 나타내고 있다.

결론적으로 두 모델 각각의 전체정확도가 모두 80%를 넘지 못하므로, 높은 수준의 정확도 수준을 갖고 있지 않다. 또한 양성(benign)에서의 정확도 평균은 90.3%, 악성(malignant)에서의 정확도 평균은 46.2%으로 양성과 악성 사이의 정확도 차이가 크게 나고있다.

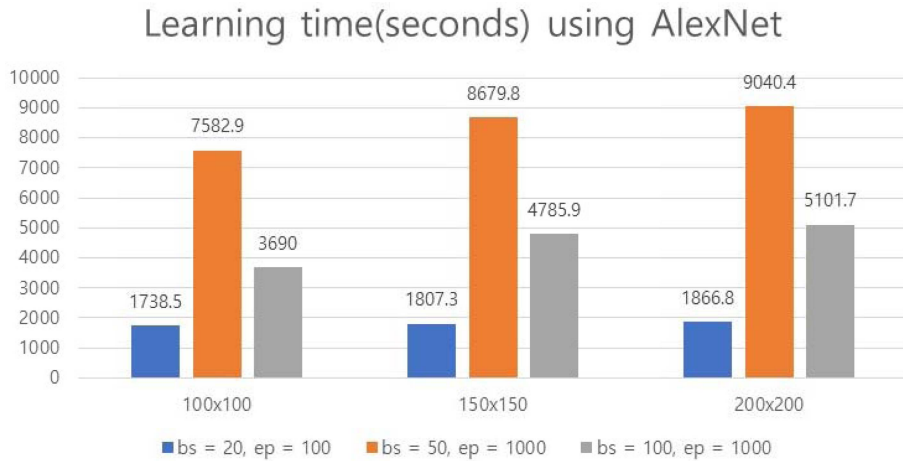


그림 4. AlexNet을 이용한 학습시간 결과.
Fig. 4. Learning time(s) using Alexnet.

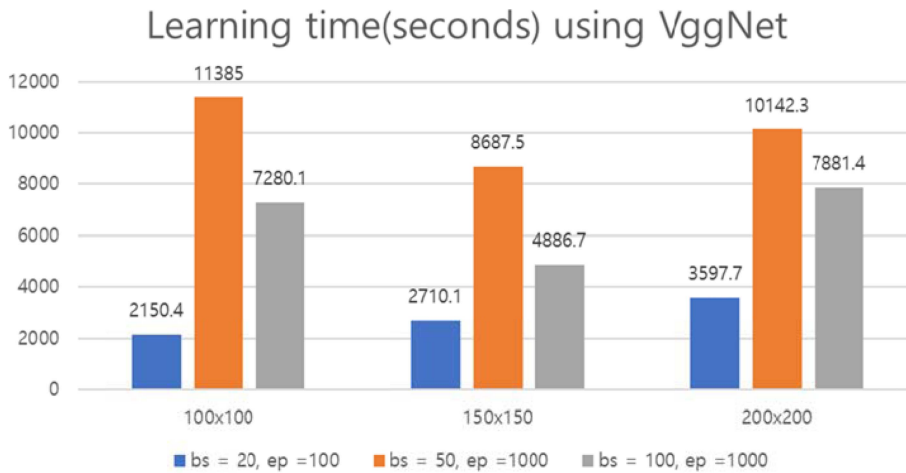


그림 5. VggNet을 이용한 학습시간 결과.
Fig 5. Learning time(s) using VggNet.

2. 이미지 크기 및 배치사이즈와 반복횟수에 차이에 따른 학습시간 분석을 통한 AlexNet과 VggNet결과 비교 및 분석

표2는 각 이미지 크기 및 배치 사이즈와 반복횟수에 따라 AlexNet과 VggNet이 양성 종양과 악성종양을 분류하는 데 걸리는 총 시간을 비교한 것이다. 그림4와 그림5는 각 이미지 크기 및 배치 사이즈와 반복횟수에 따라 AlexNet과 VggNet에서의 학습시간을 각각 나타낸 표이다. 각 이미지 크기에 따라 나타낸 3가지의 학습시간이 배치 사이즈와 반복횟수에 따라 거의 비슷하게 나타났다. VggNet에서 이미지 크기가 200일 때의 학습시간만 배치사이즈와 반복횟수에 따라 계속 증가함을 나타냈다.

딥러닝의 두 모델에서의 정확도(Accuracy)의 비는 평균적으로 1.02이므로 큰 차이가 없다. 또한 두 모델에서의 학습시간은 AlexNet이 VggNet보다 약 1.64배 빠르지만 통

계적으로 차이는 없었다($p > 0.05$).

V. 결 론

본 연구에서는 기존의 컴퓨터 보조진단과 관련하여 딥러닝을 이용한 분류방식을 연구하였다. 딥러닝에 의한 방식이 아직 편리하지는 않지만 분류의 힘은 높게 평가되었다. 아직 인공지능을 활용한 의료영상처리 방법에서 해결되어야 할 문제들이 있다. 16비트 영상에서의 분류 방식이 아직 미진하여, 본 연구에서는 8비트 영상으로 변환하여 시도되었다. 윈도우 폭과 윈도우 길이를 조절하는 방식에 대해서도 파라미터들에 대한 반응을 하는 것이 필요하다. 또한 영상에서 높이와 길이를 같은 크기로 맞추어 하는 제약이 있어서 모든 종양들에 대한 결과를 일정하게 반영하는 것도 필요하다.

본 연구에 쓰인 영상에서 환자마다 종양의 크기와 모양이 다르기 때문에 자동적으로 한번에 관심영역을 검출하는 데 있어서 한계가 있었다. 특히, 병변의 영역을 일부 결정하기 어려운 부분에 있어서는 전문의의 도움이 필요하기 때문에 여러 차례 도움을 받아 진행하였다. 그래서 원본이미지에서 관심영역의 마진을 수동적으로 결정하여 검출하는 과정에 있어서 시간이 오래 걸렸다. 이러한 수동적 방식에 있어서 영상 분할 관련 연구의 필요성이 느껴져 추후 관심영역 검출의 자동화에 대한 추가 연구를 진행하고자 한다.

원본영상에서 관심영역을 정사각형으로 잘라낸 이미지의 크기가 종양의 크기에 따라 제각기 다르다. 따라서, 이미지의 크기를 모두 일정한 크기로 변환(resampling)을 하게 되면 학습의 결과도 달라질 것이라고 추측했으나 결과에 큰 차이가 없었다. 이를 통해, CT, MRI 등을 포함한 다양한 영상에서 이미지의 크기가 매우 크거나 작더라도 일정 크기로 사이즈를 변환해도 결과에 영향을 미치지 않을 것이다. 또한 이미지의 크기와 학습시간은 비례하기 때문에 이미지의 크기를 적당한 크기로 축소하면 학습 시간을 절약할 수 있다.

본 연구에서 반복횟수와 배치사이즈 값에 따른 두 개의 분류 모델(AlexNet, VggNet)을 활용하여 결과를 획득하였다. 먼저, 반복횟수와 배치사이즈를 조금씩 다르게 두어 결과를 얻었다. 반복횟수와 배치사이즈를 각각 적게 설정하거나 크게 설정하여도 결과가 거의 비슷하고 반복횟수와 배치사이즈가 클수록 정확도가 높아지는 것이 아니다. 본 연구에서는 배치사이즈를 100으로 설정하고 반복횟수를 1000으로 했을 경우 가장 적합하다고 생각된다. 하지만 데이터셋(Data Set)의 양에 따라 배치사이즈와 반복횟수는 다르게 설정해야 할 것이다.

AlexNet과 VggNet으로 학습시킨 결과를 표1을 통해 보면 양성(benign)과 악성(malignant)사이의 정확도차이가 두드러지게 나타났음을 알 수 있다. 새로운 환자가 인공지능을 통해 진단을 받을 때, 양성의 정확도보다 악성의 정확도가 더 높게 나타나야 안전하다고 할 수 있는데 이 실험에서는 반대의 결과가 나타났다. 추후 악성의 정확도를 더 높이기 위해 연구하고자 한다.

본 연구에서는 각 이미지 크기별로 AlexNet과 VggNet이 양성종양과 악성종양을 분류하는 데 걸리는 총 시간을 비교하였다. 분류하는 데 걸리는 학습 시간을 통해 2가지 사실을 알 수 있다.

첫째는 배치사이즈와 반복횟수의 값의 차이가 클수록 더 많은 시간이 걸린다. 표2에서 배치사이즈가 50이고 반복횟수가 1000일때가 배치사이즈가 100이고 반복횟수가 1000일 때보다 2배에 가까운 시간이 더 소요된다. 결과적으로 반복횟수와 배치 사이즈 값의 차이를 크게 둘수록 학습하는

데 걸리는 시간이 오래 걸리기 때문에 반복횟수와 배치사이즈 사이의 값을 크게 두지 않으면 시간을 절약할 수 있다.

두번째는 똑같은 조건에서 두가지 모델을 활용하여 학습을 시킬 때, AlexNet의 정확도가 VggNet의 정확도보다 조금 우세하였고 똑같은 조건에서 두 모델을 각각 돌렸을 때 AlexNet이 VggNet보다 속도가 항상 빠르다는 것이다. 두 모델의 정확도(Accuracy)의 차이는 크게 나타나지 않고 비슷하지만 평균적으로 AlexNet의 정확도가 약간 우세하였다. 또한 학습시간에서 AlexNet이 VggNet보다 최대 1.6배정도 빠른 속도로 학습한다. 두 모델 모두 전체적인 정확도 부분에서는 90%에 많이 미치지 못하기 때문에 새로운 모델을 활용하여 간 종양 데이터를 학습시키거나, 데이터 영상을 전처리하는 과정을 하는 등 좀 더 많은 연구 개발이 필요하다고 생각된다. 하지만 본 연구에서 AlexNet과 VggNet의 정확도와 학습시간을 비교했을 때, AlexNet이 VggNet보다 성능이 좀 더 우수하다고 생각되며 분류모델로 AlexNet을 활용하면 효율적이고 더 나은 결과를 얻을 수 있을 것이라 생각된다.

최종적으로, 본 연구를 통해 이미지 크기가 학습시간과 비례한다는 결과를 통해 앞으로 수많은 데이터 이미지들을 활용한 딥러닝 학습을 시킬 때에 적당한 이미지 크기를 선택함으로써, 시간을 절약하고 메모리 양을 크게 줄이는 데에 많은 도움을 줄 수 있을 것에 기대된다. 또한 두 모델의 정확도와 학습시간을 비교함으로써 딥러닝을 적용하려는 여러 연구들에게 파라미터에 따른 정보 제공에 도움이 될 것이며 모델을 선택하는 데에도 도움이 될 것이라고 생각된다.

참고문헌

- [1] J. Shim, "Current Status of Research on Liver Cancer in Korea," *The Korean Liver Cancer Journal*, vol. 17, no. 2, pp.105-112, 2017.
- [2] <https://www.cancer.go.kr/lay1/SIT639C640/contents.do>, National Cancer Information Center, accessed on June. 13, 2018
- [3] M.A. Lee, "Special Review: Screening and surveillance of hepatocellular carcinoma," *Korean Journal of Medicine*, vol. 79, no. 3, pp.224-230, 2010.
- [4] T. S.Furey, N. Cristianini, N. Duffy, D. W.Bendnarski, M. Schummer and D. Haussler, "Support vector machine classification and validation of cancer tissue samples using microarray expression data," *Bioinformatics*, vol. 16, no. 10, pp. 906-914, 2000.
- [5] M. Långkvist, A. Kiselev, M. Alirezaie and A. Loutfi, "Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks," *Remote Sensing*, vol. 8, no. 4, pp. 329, 2016.
- [6] J. Zhang, K.K. Ma, M.H. Er and V. Chong, "Tumor Segmentation from Magnetic Resonance Imaging by Learning via one-class support vector machine," in *International Work-*

- shop on *Advanced Image Technology (IWAIT'04)*, Singapore, Jan 2004, pp. 207-211.
- [7] S. Zhang, L. Yang, C. Peng and M. Wu, "Logistic regression analysis of risk factors for postoperative recurrence of spinal tumors and analysis of prognostic factors," *Oncology Letters*, vol. 15, no. 2, pp. 1716-1722, 2018.
- [8] H.P. Chan, B. Sahiner, N. Petrick, M.A. Helvie, K.L. Lam, D.D. Adler and M.M. Goodsitt, "Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network," *Physics in Medicine & Biology*, vol. 42, no. 3, pp. 549, 1997.
- [9] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega and D. R. Webster, "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," *JAMA*, vol. 316, no. 22, pp. 2402-2410, 2016.
- [10] H.C. Park, E.J. Lee, Y.J. Kim, S.W. Lee, "Colonoscopy Image Classification Using Convolutional Neural Networks," *The Korean Institute of Next Generation Computing*, vol. 14, no. 1, pp. 89-98, 2018.
- [11] H.W. Kim, B.N. Kim, J.E. Lee, W.S. Jang and S.K. Yoo, "Performance Improvement of Convolutional Neural Network for Pulmonary Nodule Detection," *Journal of Biomedical Engineering Research*, vol. 38, no. 5, pp. 237-241, 2017.
- [12] J. H. Lee, H. S. Ahn, D. H. Choi, and K. S. Tae, "Evaluation on the Usefulness of X-ray Computer-Aided Detection (CAD) System for Pulmonary Tuberculosis (PTB) using SegNet", *Journal of Biomedical Engineering Research*, vol. 38, no.1, pp. 25-31, 2017
- [13] L. Gattinoni, P. Caironi, P. Pelosi and L. R. Goodman, "What Has Computed Tomography Taught Us about the," *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE*, vol. 164, no. 9, pp. 1701-1711, 2001.
- [14] S. Sabouri., A. Khatami, P. Azadeh, J. Ghoroubi and Gh. Azimi, "Adding Liver Window Setting to the Standard Abdominal CT Scan Protocol: Is It Useful?," *Iranian Journal of Radiology*, vol. 5, no. 2, pp. 65-70, 2008.
- [15] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *In Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *arXiv preprint arXiv:1409.1556*, Apr. 2014.