

심층 컨볼루션 신경망을 이용한 OCT 볼륨 데이터로부터 AMD 진단

권오흠[†], 정유진^{**}, 송하주^{***}

AMD Identification from OCT Volume Data using Deep Convolutional Neural Network

Oh-Heum Kwon[†], Yoo Jin Jung^{**}, Ha-Joo Song^{***}

ABSTRACT

Optical coherence tomography (OCT) is the most common medical imaging device with the ability to image the retina in eyes at micrometer resolution and to visualize the pathological indicators of many retinal diseases such as Age-Related Macular Degeneration (AMD) and diabetic retinopathy. Accordingly, there have been research activities to analyze and process OCT images to identify those indicators and make medical decisions based on the findings. In this paper, we use a deep convolutional neural network for analysis of OCT volume data to distinguish AMD from normal patients. We propose a novel approach where images in each OCT volume are grouped together into sub-volumes and the network is trained by those sub-volumes instead of individual images. We conducted an experiment using public data set to evaluate the performance of the proposed approach. The experiment confirmed performance improvement of our approach over the traditional image-by-image training approach.

Key words: Optical Coherence Tomography, Age-Related Macular Degeneration, Deep Learning, Convolutional Neural Network

1. 서 론

OCT (Optical Coherence Tomography)는 적외선 간섭측정(interferometry)을 이용하여 인체 조직을 침투하는 이미지를 획득함으로써 조직의 표면 아래의 해부학적 구조를 시각화하는 의료 영상 기기이다. 안과, 조직 병리학, 피부 암 진단 등의 다양한 의료 분야에서 질병의 진단과 치료에 광범위하게 사용되고 있다. 특히 안구의 후면인 망막(retina)을 수천 분의 일 밀리미터 수준의 해상도로 시각화할 수 있으

로, 망막 층(layer)의 시각화가 가능하여 다양한 안과 병리학적 지표에 대한 진단 장비로 사용되고 있다. 치료를 시작하기 전에 망막의 상태를 평가하고 치료의 과정에서 효과를 관찰하기 위한 가장 중요하며 가장 자주 사용되는 장비이다. 미국의 경우 매년 4~5백만 명이 OCT 진단을 사용하며 그 빈도는 급속하게 늘어나고 있는 것으로 알려져 있다[1]. Fig. 1은 망막에 대한 OCT 이미지의 예이다.

OCT가 개발되고 의료 현장에 적용되기 시작한 이후로 컴퓨터 비전(vision) 및 이미지 처리 기술을

* Corresponding Author: Oh-Heum Kwon, Address: (48513) Yongso-ro 45, Nam-gu, Busan, Korea, TEL: +82-51-629-6255, FAX: +82-51-629-6264, E-mail: ohkwn@pknu.ac.kr

Receipt date: Jul. 17, 2017, Approval date: Aug. 1, 2017
[†] Dept. of IT Convergence and Application, Pukyong National University

^{**} Division of Computer & Electronic Systems Engineering, Hankuk University of Foreign Studies (E-mail: chungyj@hufs.ac.kr)

^{***} Dept. of IT Convergence and Application, Pukyong National University (E-mail: hajoosong@pknu.ac.kr)

* This research was supported by Pukyong National University Research Grant CD20160563

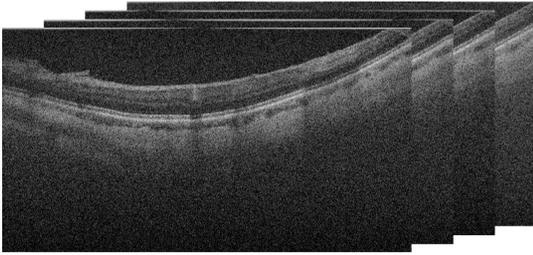


Fig. 1. OCT image examples.

OCT 이미지의 분석과 처리에 적용하고자 하는 지속적인 연구가 있어 왔다. 대표적인 예로는 OCT 이미지로부터 망막 층이나 수액(fluid) 영역을 자동으로 인식하고 분리하는(segmentation) 기법에 관한 것이 있다 [2]. 이러한 층의 두께는 일부 망막 질환의 유무를 판정하는 지표로 알려져 있다. 또한 OCT 이미지로부터 각막과 망막 사이의 맥락막(choroid) 영역의 무늬(texture)나 2차원 혹은 3차원의 형태적(morphological) 특징을 추출하고 이를 안과적 질병의 진단에 활용하는 다양한 연구들이 진행되고 있다 [3,4].

한편 최근 기계학습(machine learning) 분야의 급속한 발전과 더불어 기계 학습 기술을 의료 이미지에 적용하여 자동진단을 하는 것에 관한 연구가 활발하게 이루어지고 있다. 자동 진단은 의료 전문가의 진단 과정의 효율성을 높이고 또한 원격 진료 등의 가능성을 높이는 매우 중요한 연구 분야이다. 특히 심층(deep) 컨볼루션 신경망(convolutional neural network: CNN)은 이미지의 지역적인 특징을 추출하기 위한 여러 겹의 컨볼루션(convolution) 층을 사용하는 심층 신경망의 한 종류로서 대용량의 이미지 및 비디오 인식에서 매우 효과적임이 입증되고 있으며 다양한 응용 분야에서 성공적으로 활용되고 있다 [5, 6,7]. 심층 CNN을 사용하여 OCT 이미지로부터 녹내장(glaucoma), 노년기황반변성 (Age-Related Macular Degeneration: AMD), 당뇨 망막증(diabetic retinopathy) 등의 다양한 안과 질환을 자동으로 진단하는 것에 관한 연구가 활발하게 진행되고 있다. 예를 들어 Abramoff 등은 심층 CNN을 이용하여 874명의 환자에 대한 OCT 이미지로부터 당뇨 망막증을 자동으로 식별하는 연구를 수행하였다 [8]. Asaoka 등은 역시 CNN을 이용하여 녹내장을 자동 진단하는 연구를 수행하였다 [9].

본 논문은 심층 CNN을 이용하여 OCT 이미지로

부터 대표적인 망막 질환의 하나인 AMD 즉 노년기 황반변성을 자동 진단하는 문제를 다룬다. AMD는 세계적으로 대략 1억 6천만 명 이상의 인구가 앓고 있는 망막 질환으로 노인성 실명의 대표적인 원인 중 하나이다. 유전적 요인과 흡연 등의 생활 습관, 그리고 환경적 요인이 주요 병인으로 알려져 있다 [10]. 정상적인 망막과 구분되는 AMD 증상들 중에서 망막 색소 상피(retinal pigment epithelium)의 변화, 망막 내 수액(intraretinal fluid), 결정체(drusen) 등이 OCT 이미지로부터 식별될 수 있는 시각적인 특징들로 알려져 있다.

CNN을 이용하여 망막에 대한 이러한 시각적 특징을 식별함으로써 AMD 자동 진단을 수행하는 것을 목적으로 하는 연구들이 최근에 활발하게 진행되고 있다. Schlegl 등은 OCT 이미지 내에서 병리적 이상 영역의 위치에 대한 전문가의 추가 정보가 제공된다는 가정 하에 OCT 이미지로부터 망막 내 낭종(intraretinal cysts: IRC)과 정상 영역을 분리하는 문제에 대해서 CNN을 사용하였다 [11]. Cecilia 등은 OCT 볼륨(volume)을 구성하는 개별 이미지에 대해서 신경망을 이용해 AMD 여부를 판별하고, 그 결과를 평균하여 볼륨 혹은 환자 단위의 진단을 내리는 방식을 사용하였으며, 개별 이미지 단위로는 87.63%, OCT 볼륨 단위로는 88.98%의 정확도(accuracy)를 달성함으로써 CNN이 AMD 자동진단을 위한 효과적인 수단이 될 수 있음을 보였다[12]. Apostolopoulos 등은 2단계의 CNN을 통해서 OCT 볼륨으로부터 AMD를 자동 진단하는 방법을 제시하였다[10]. 1단계에 해당하는 RetiNet B는 개별 이미지 단위로 AMD 여부를 판정하는 신경망이다. 2단계인 RetiNet C는 미리 학습된(pretrained) RetiNet B를 전방 층으로 가지고, 전방 층의 출력으로부터 볼륨 단위의 판정을 내리는 완전 연결된(fully-connected) 후방 층이 추가된 구조를 가진다.

OCT 이미지로부터 AMD를 자동 진단하는 연구에서의 한 가지 기본적인 문제점은 OCT 이미지들이 볼륨 단위로 라벨링(labeling)되어 있다는 것이다. 한번의 OCT 촬영으로 기기 혹은 설정에 따라서 적게는 50개에서 많게는 1000개 이상의 Fig. 1에서 예시한 것과 같은 단면(cross-sectional) 이미지들이 촬영되고, 의사는 이 이미지들을 종합적으로 관찰하여 AMD 여부에 대한 진단을 내리게 된다. AMD 환자

라고 하더라도 하나의 볼륨을 구성하는 모든 이미지들에서 해당 증상이 관찰되지는 않으며, 특히 질병의 초기 단계인 경우에는 진단을 위해서 여러 장의 연속된 단면 이미지를 동시에 관찰하면서 파악하기 힘든 희미한 징후를 찾아내야하기도 한다. 또한 촬영 과정에서 환자의 미세한 움직임 등으로 인하여 품질이 저하된 이미지가 포함되기도 한다. 결과적으로 질병 유무에 대한 판정은 개별 이미지 단위가 아니라 볼륨 전체에 대해서 내려지게 되며, 따라서 신경망의 학습에 사용하기 위한 데이터들은 기본적으로 볼륨 단위의 라벨을 가지게 된다.

따라서 엄밀한 의미에서 지도 학습(supervised learning)이 이루어지기 위해서는 각각의 OCT 볼륨이 하나의 데이터 항목으로 간주되어야 할 것이다. 하지만 현실적으로 대량의 의료 데이터를 수집하기 어려운 측면이 있으므로 이렇게 각각의 OCT 볼륨을 하나의 데이터로 취급하는 방법은 학습 데이터의 부족에 직면할 수밖에 없다. 그래서 대부분의 기존 연구에서는 각 이미지 별로 라벨을 수작업으로 추가하거나, 하나의 OCT 볼륨으로 부터 10장 내외의 작은 개수의 이미지만을 선택하여 연구에 사용하기도 하고[12], 혹은 잘못된 라벨의 위험을 감수하고 볼륨 라벨을 그냥 이미지 라벨로 간주하는 접근을 취하게 된다[10]. 전자의 경우에는 수 만장 혹은 그 이상의 이미지 각각을 전문가가 판독하는 과정을 거쳐야 하는데 이는 많은 시간과 비용을 요구하는 작업이다. 후자의 경우에는 일부의 이미지에 대해서 잘못된 라벨을 부여하는 것을 감수한다는 의미에서 일종의 약한 지도 학습(weak supervised learning)이라고 할 수 있다.

본 논문에서는 이러한 기본적인 문제를 완화하기 위해서 다음과 같은 접근법을 사용한다. 우리는 각각의 OCT 볼륨 데이터로부터 복수개의 연속된 이미지 프레임을 선택하여 이것을 하나의 서브 볼륨(sub-volume)이라고 부르고, 각각의 서브 볼륨을 신경망의 학습과 테스트를 위한 데이터 항목으로 사용한다. 각각의 서브 볼륨에게는 자신이 속한 볼륨의 라벨을 부여한다. 이렇게 서브 볼륨 단위로 학습과 판정을 하는 것은 데이터에 잘못된 라벨을 부여함으로써 잘못된 학습이 일어날 가능성을 줄이는데 도움이 된다. 즉 서브 볼륨을 구성하는 이미지 중에서 적어도 하나의 이미지에 질병의 증상이 존재할 확률은 서브 볼륨

의 크기가 증가함에 따라 증가하므로 개별 이미지에 대해서 볼륨 라벨을 부여하는 것보다 잘못된 라벨이 부여될 확률이 감소하게 될 것이다. 또한 한 장의 이미지에 온전하게 드러나지 않고 연속된 여러 장의 이미지를 통해서 관찰할 수 있는 불분명한 징후를 파악하는데 도움이 될 수 있다. 즉 원래의 볼륨이 내포한 3차원적 정보가 어느 정도 활용되는 것을 기대해 볼 수 있다. 서브 볼륨들은 슬라이딩 윈도우(sliding window) 방식으로 서로 중첩되게 구성하므로 학습을 위한 데이터의 개수 역시 개별 이미지 단위로 하는 방법에 비해 크게 줄어들지 않는다. 테스트 볼륨에 대해서는 이렇게 서브 볼륨 단위로 계산된 확률의 평균을 이용하여 볼륨 단위의 판정을 내리는 기본적인 접근을 택하였다.

만약 서브 볼륨의 크기가 1이면 이는 기존의 연구들과 동일하게 이미지 단위의 라벨을 사용하는 방법으로 환원된다. 본 논문에서는 서브 볼륨 단위로 입력을 받는 깊은 컨볼루션 신경망을 구성하고 개별 이미지를 사용하는 경우와 크기가 3인 서브 볼륨을 사용하는 경우에 대해서 진단의 정확성(accuracy)과 민감도(sensitivity)에 대한 비교 분석을 수행하였다. 분석의 결과 크기가 3인 서브 볼륨을 사용하는 것이 개별 이미지 단위로 학습하는 경우보다 향상된 성능을 보여주었다. 이미지 혹은 서브 볼륨 단위의 진단 결과를 평균하여 볼륨 단위의 진단을 했을 때도 역시 서브 볼륨 단위의 인식이 개별 이미지를 사용했을 때 보다 우위를 보였다. 이는 본 논문이 제시하는 서브 볼륨 단위의 학습과 인식이 효과적인 방법이 될 수 있음을 입증하는 결과이다.

또한 AMD로 진단된 이미지에서 그러한 진단을 내리는데 가장 크게 기여한 이미지 영역을 추출하기 위해서 차단(occlusion) 검사를 수행하였다. 차단 검사는 이미지의 모든 가능한 위치에 대해서 적절한 크기의 마스크 윈도우를 이용하여 원래 이미지를 가린 후 검사를 하고 그 결과를 원래 이미지에 대한 진단 결과와 비교함으로써 이미지의 각 부분이 진단에 기여한 정도를 분석하는 것이다. 테스트 결과 신경망이 OCT 이미지 내에서 병리적 증상 부위를 적절하게 추출하는 것을 확인할 수 있었다.

본 논문에서는 OCT 이미지 볼륨으로부터 AMD를 자동 진단하는 심층 CNN을 제시하고, 서브 볼륨 단위의 인식이 개별 이미지 단위의 인식에 비해 효과

적임을 실험을 통하여 보였다. 본 논문의 구성은 다음과 같다. 우선 2장에서는 연구 방법과 본 연구에서 사용한 신경망에 대해서 기술한다. 3장에서는 성능 평가의 방법과 결과를 기술한다. 4장에서는 결론을 제시한다.

2. 연구방법

2.1 용어와 표기

OCT 데이터의 집합은 n 개의 볼륨 V_1, V_2, \dots, V_n 으로 구성된다. $V_i, i=1, \dots, n_1$,은 AMD 환자로부터 획득한 OCT 데이터이며, $V_j, j=n_1+1, \dots, n$,은 정상 환자의 것이다. 각각의 볼륨 V_i 는 K 개의 동일한 크기의 1채널 이미지로 구성된다. 볼륨 V_i 를 구성하는 이미지들을 $I_{i,1}, I_{i,2}, \dots, I_{i,K}$ 로 표시한다. 이 이미지들은 공간적으로 연속된 단면 이미지들이다. 각각의 볼륨 V_i 는 대응하는 라벨(label) $Y_i \in \{0, 1\}$ 을 가진다. 라벨은 볼륨 단위로 존재하며 각각의 이미지에 대응하는 라벨은 제공되지 않는다.

하나의 볼륨으로부터 k 개의 연속된 이미지를 선택하여 서브 볼륨을 구성하고, 서브 볼륨에는 그것이 속한 볼륨의 라벨을 부여한다. 볼륨 V_i 로부터 추출한 서브 볼륨들을 $S_{ij}, j=1, \dots, K-k+1$,로 표시한다. 서브 볼륨 S_{ij} 는 볼륨 V_i 에 속한 연속된 이미지들 ($I_j, I_{j+1}, \dots, I_{j+k-1}$)로 구성된다. 서브 볼륨들은 이렇게 슬라이딩 윈도우 방식으로 서로 중첩되도록 구성한다. 볼륨의 가장자리 이미지들, 즉 $I_{i,1}, \dots, I_{i,k-1}$ 과 $I_{i,k-k+2}, \dots, I_{i,K}$ 는 나머지 이미지들보다 더 적은 개수의 서브 볼륨에 속하게 되므로 가장자리 이미지들이 볼륨의 중심 부분의 이미지들 보다 상대적으로 덜 자주 학습에 이용되는 부수적인 효과가 있다. 서브 볼륨의 크기 k 가 클수록 서브 볼륨 라벨은 진실과 일치할 확률이 높아지고, 만약 $k=1$ 이라면 각각의 서브 볼륨은 단일 이미지로 구성되므로 기존의 연구들과 유사한 이미지 단위의 판별이 된다.

신경망은 출력 층에서 각각의 서브 볼륨에 대해서 AMD일 확률을 제공한다. 볼륨 전체에 대한 판정은 볼륨에 속한 서브 볼륨들의 확률의 평균을 계산하여 내린다. 즉 볼륨 V_i 가 AMD일 확률은 $\frac{1}{K-k+1} \sum_{j=1}^{K-k+1} p_{ij}$ 이고, 여기서 p_{ij} 는 서브볼륨 S_{ij} 가 AMD일 확률이다.

2.2 신경망 구성

컨볼루션 신경망이 이미지 및 비디오 인식 분야에서 커다란 성공을 성취하는 과정에서 네트워크의 구조와 그것을 구성하는 세부 구성요소에 대한 지속적인 개선과 발전이 이루어지고 있고 그 결과들은 많은 응용 연구들에 의해서 공유되고 있다. 이 연구에서 우리는 새로운 신경망 구조를 설계하는 대신 최근에 제안된 이미지 인식을 위한 신경망 구조들 중 가장 성공적인 사례의 하나인 VGG19 네트워크를 우리의 설정과 상황에 맞게 부분적으로 수정하여 사용하였다 [13]. 사용된 네트워크는 Fig. 2에서 설명하는 바와 같이 우선 크게 전방을 구성하는 5 그룹의 컨볼루션 층들과 후방을 구성하는 3겹의 완전연결(fully-connected: FC) 층들로 구성되는 매우 깊은 신경망이다. 컨볼루션 층 그룹들은 각각 다시 2~4개의 연속된 컨볼루션 층으로 구성되며 각 그룹의 마지막에는 맥스풀링(maxpooling) 층이 추가된다. 각각의 컨볼루션 층에서는 매우 작은 크기의 3×3 필터들이 적용되고, 스트라이드(stride) 크기는 1 혹은 2가 사용된다. FC층에서 최종 출력은 softmax 연산을 적용하여 AMD일 확률의 형태로 출력된다. 또한 모든 은닉(hidden) 층에는 ReLU (rectification) 비선형화 함수가 사용된다.

입력 층에서는 각각의 서브 볼륨이 하나의 k -채널 이미지의 형태로 공급된다. 서브 볼륨은 입력되기 전에 전체 데이터의 평균 픽셀 밝기를 빼는 통상적인 전처리 작업을 적용하였으며, 개별 이미지의 사이즈는 256×512이다. VGG19과 비교하여 입력 층의 크기가 입력 이미지의 크기에 맞게 변경되었으며, 마지막 FC3 층이 추가된 형태이다. 트레이닝은 평균 제곱근(mean square) 목적 함수(objective function)와 경사 하강법(gradient descent)을 사용하여 이루어졌다. 트레이닝 시간을 단축하기 위해서 네트워크의 첫

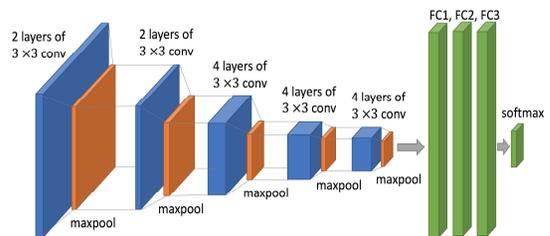


Fig. 2. CNN architecture.

번째 컨볼루션 층과 FC1 및 FC3 층을 제외한 나머지 층들은 미리 트레이닝 된(pre-trained) 가중값을 이용하여 초기화한 후 트레이닝을 지속하였다. Fig. 2는 사용된 네트워크의 구조를 요약한 것이다.

2.3 학습과 테스트

학습과 테스트를 위한 데이터의 부족을 보완하기 위해서 전체 데이터 셋을 5개의 동일 크기의 집합으로 랜덤하게 분할한 후 그 중 하나를 테스트 데이터로 사용하고 나머지 4개를 학습 데이터로 사용하는 5겹 교차 검증(5-fold cross validation)을 수행하였다. 즉 5개의 동일한 신경망을 구성하고 각각에 대해서 따로 학습과 테스트를 실시한다. 각각의 네트워크는 5분할된 데이터 집합들 중 하나를 테스트 데이터로 사용되고 나머지 4개의 데이터 집합을 학습 데이터로 사용된다. 물론 학습의 과정에서 네트워크 간에는 어떤 정보도 교환되거나 공유되지 않는다. 네트워크의 성능은 5개의 네트워크의 결과를 평균하여 평가한다.

학습 데이터가 m_1 개의 AMD 볼륨과 m_2 개의 정상 볼륨으로 구성되고, 학습 배치(training batch)의 크기를 B 라고 할 때 모든 데이터 들이 동일한 세대(epoch) 동안 학습되도록 하기 위해서 하나의 배치는 $Bm_1/(m_1+m_2)$ 개의 AMD 서브 볼륨과 $Bm_2/(m_1+m_2)$ 개의 정상 서브 볼륨으로 구성한다. 서로 다른 볼륨 내에서 유사한 위치에 있는 이미지들은 망막의 유사한 위치에 대한 이미지들이므로 서로 일정한 공통성을 가진다. 하나의 학습 배치가 이렇게 위치적으로 동질성을 가진 서브 볼륨들로만 구성되는 것을 방지하기 위해서 각 학습 볼륨에 속한 $K-k+1$ 개의 서브 볼륨들은 랜덤 순서로 배열한 후 그 순서대로 추출하여 학습 배치에 소속시킨다. 학습 과정에서 한 세대가 종료될 때마다 각 볼륨 내의 서브 볼륨들은 다시 랜덤 순서로 재배열 한다. 이렇게 하여 학습 기간 동안 가능한 한 다양한 내용의 배치가 구성되도록 하였다.

3. 성능 평가

3.1 데이터 및 실험 환경

본 논문에서는 Duke 대학의 VIP (Vision and Image Processing) 연구실에서 AMD 환자의 OCT 이미지에 대한 분석을 목적으로 제공하는 공개 데이

터 셋을 사용하였다 [14]. 본 연구 데이터 셋은 총 384개의 OCT 볼륨으로 구성된다. 그 중 269개는 AMD 환자의 것이고, 나머지 115개는 건강한 사람으로부터 획득한 것이다. 각각의 OCT 볼륨은 100개의 순차적인 이미지로 구성되고, 각 이미지의 해상도는 500×1000 이다. 이 이미지들을 256×512 크기로 축소하여 사용하였다. 서브 볼륨의 크기는 $k=1$ 인 경우와 $k=3$ 인 두 가지 경우에 대해서 실험을 수행하였다. 이하에서는 $k=1$ 인 경우를 “이미지 단위의 학습”, $k=3$ 인 경우를 “서브 볼륨 단위의 학습”이라고 부르겠다. 학습 배치의 크기는 이미지 단위인 경우에는 60으로, 서브 볼륨 단위인 경우에는 20으로 하여 하나의 배치에 포함된 이미지의 총 개수가 60개로 동일하도록 유지 하였다. 신경망은 Google의 Tensorflow 버전 1.0을 이용하여 구현하였으며 한 개의 Nvidia 1080Ti GPU를 사용하는 Ubuntu16.04 시스템에서 실험을 수행하였다.

3.2 실험 결과

학습은 10,000 단계 동안 반복하였으며 이는 대략 20 세대(epoch) 정도에 해당한다. Fig. 3은 서브 볼륨 단위의 학습과정에서 트레이닝 손실(loss)과 정확도(accuracy)의 변화과정의 한 예를 보여준다. 대략 8,000에서 10,000 단계를 전후하여 더 이상 학습이 진행되지 않는 것을 볼 수 있다. 이는 원래의 VGG19 네트워크의 트레이닝 시간과 비교하여 매우 빠른 진행인데 신경망의 중간 층에서 미리 학습된 가중치를 초기값으로 사용했기 때문으로 보인다.

이미지 단위의 학습과 서브 볼륨 단위의 학습에 대해서 진단의 정확성을 평가하기 위해 5겹 교차검증의 결과를 모두 합산하여 FPR (False Positive Rate)와 FNR(False Negative Rate)를 분석하였다. 이미지 단위의 학습의 경우 각각의 이미지에 대한 FPR는 10.13%, FNR는 4.34%였다. 반면 서브 볼륨 단위의 경우 각각의 서브 볼륨에 대한 FPR은 8.43%, FNR은 5.25%였다. 각 볼륨에 속한 모든 이미지 혹은 서브 볼륨에 대한 진단 결과를 평균하여 볼륨 단위의 판정을 한 경우에 대한 정확도 분석의 결과는 다음과 같다. 이미지 단위 학습의 결과 FPR은 7.01%이고 FNR는 1.48%였다. 이는 115개의 정상 볼륨과 269명의 AMD 볼륨에 대해서 각각 8개와 4개의 볼륨에 대해서 잘못된 판정을 내린 것이다. 서브 볼륨 단위

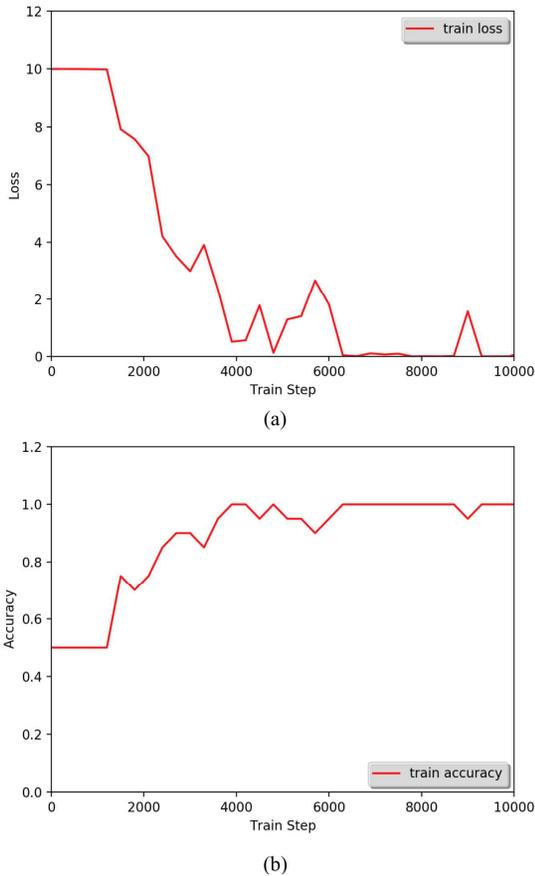


Fig. 3. Training loss and accuracy.

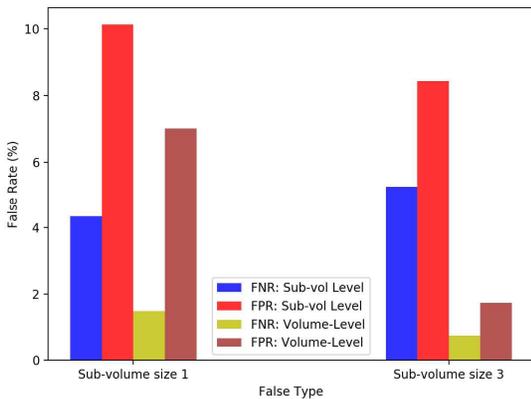


Fig. 4. Comparison of FNR and FPR.

의 학습에서는 FPR이 1.75%, 그리고 FNR은 0.74%였다. 이는 2개의 정상볼륨을 AMD 볼륨으로, 2개의 AMD 볼륨을 정상 볼륨으로 잘못 판정한 결과이다. Fig 4는 이 결과를 요약한 것이다. 주목할 점은 모든

경우에 대해서 정상 이미지를 AMD로 잘못 판정하는 오류가 그 반대의 경우보다 높았다는 것이다. 이것이 네트워크의 구조나 학습 기법의 어떤 측면에서 기인한 것인지는 현재로서는 불분명하다.

Fig. 5는 진단 임계값(threshold)의 변화에 따라서 FPR와 TPR 간의 관계를 표현하는 ROC (Receiver Operating Characteristic) 커브이다. 그림에서 점선은 서브 볼륨 수준의 판정에 대한 커브이고, 실선은 볼륨 단위의 판정에 대한 커브이다. 파란색은 크기 3인 서브 볼륨을 사용한 경우이고 빨간 색은 개별 이미지를 사용한 경우이다. AUROC(Area Under ROC)는 이미지 단위의 경우 각각 0.9855와 0.9957이고 서브 볼륨 단위의 경우 0.9879와 0.9963이었다. 이상의 실험 결과들은 일관되게 서브 볼륨 단위의 학습이 개별 이미지 단위의 학습에 비해서 일정한 성능의 향상이 있음을 보여준다.

신경망의 성능은 사용된 학습 데이터와 학습 및 테스트 과정의 다양한 세부 설정의 차이에 따라서 달라질 수 있으므로 본 연구의 결과를 기존 연구의 결과와 직접적으로 비교하여 성능의 우위를 논하는 것에는 어려움이 있다. Cecilia 등[12]의 연구는 대략 100,000장의 비공개 OCT 이미지를 사용했으며 볼륨 레벨에서 대략 90% 정도의 정확도를 달성했다고 밝히고 있다. Apostolopoulos 등[10]의 연구는 99.7%의 볼륨 레벨 AUROC를 달성하였지만 FNR과 FPR에 대해서는 구체적인 수치를 밝히지 않았고, 또한 신경망 구현의 세부적인 사항이 공개되어 있지 않아 동일한 조건의 직접적인 비교를 수행하지는 못했다. 하지만 이러한 간접적인 비교로부터 본 연구의 결과가

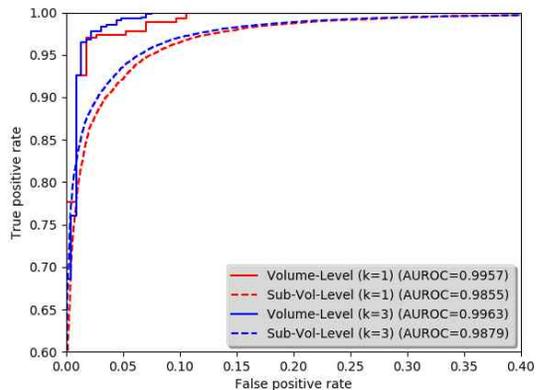


Fig. 5. ROC curves for image and sub-volume level classification.

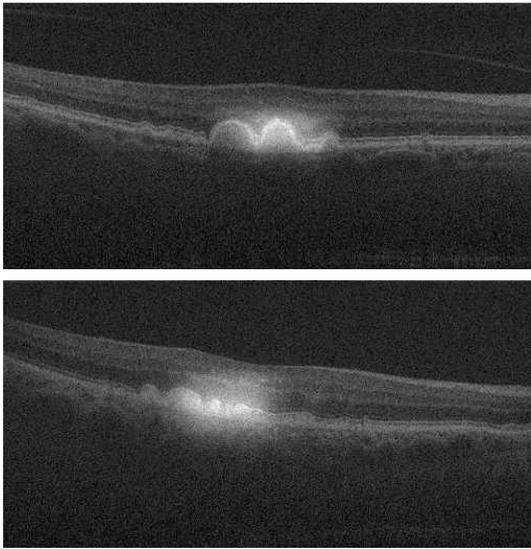


Fig. 6. Examples of occlusion test for AMD data.

기존 연구의 결과들과 충분히 비교할만하다는 것을 알 수 있다.

마지막으로 Fig. 6은 차단 검사를 실행한 결과의 예들이다. 이 검사는 모든 픽셀이 0의 값을 가지는 20×40 크기의 마스크 윈도우를 AMD로 판정된 검사 이미지의 모든 행과 열에 대해 슬라이딩 시키면서 마스크링 한 후 검사하고, 그 결과를 원래 이미지의 검사결과와 비교함으로써 이미지의 각 영역이 AMD 진단에 미치는 영향의 정도를 분석하는 것이다. 그림에서 밝게 표시된 영역일수록 AMD 진단에 더 많은 영향을 끼쳤다는 의미이다. 그림의 예들은 신경망에 의해서 병리적 이상 부위가 적절하게 식별되고 있음을 보여주는 예들이다. 실제 테스트를 수행한 대부분의 이미지들에 대해서 유사한 결과를 얻을 수 있었다.

4. 결 론

전자의무기록(Electronic Medical Record: EMR)이 보편화 되면서 디지털 이미지 처리 기술과 인공지능 기술을 의학 기술에 적용하는 사례가 급속하게 증가하고 있다. 본 논문에서는 심층 신경망 기술을 이용한 OCT 이미지 분석을 통하여 대표적인 망막질환인 AMD를 정상 환자로부터 성공적으로 식별할 수 있음을 보였다. 여러 장의 이미지로 구성되는 서브 볼륨 단위의 학습과 진단을 수행하는 접근이 기존의 개별 이미지 단위의 학습에 비해 우수한 성능을

보임을 실험을 통하여 보였다. 또한 차단 검사를 통하여 학습된 신경망이 적절하게 병리적인 이상 부위를 식별한다는 것을 보였다.

본 연구의 한계는 기본적으로 비교적 소규모의 공개 데이터에 기반하고 있다는 점이다. 보다 많은 양질의 의료 데이터를 확보하여 본 연구의 결과를 추가 검증하고 확장하는 것이 추후의 연구과제이다. 또한 AMD 이외의 녹내장 등의 다른 대표적인 안과 질환으로 연구의 영역을 넓혀가는 것 역시 향후 연구의 방향이다.

REFERENCE

- [1] J. Welzel, "Optical Coherence Tomography in Dermatology : A Review," *Skin Research and Technology*, Vol. 7, No. 1, pp. 1-9, 2001.
- [2] S.J. Chiu, J.A. Izatt, R.V. O'Connell, K.P. Winter, C.A. Toth, and S. Farsiu, "Validated Automatic Segmentation of AMD Pathology Including Drusen and Geographic Atrophy in SD-OCT Images," *Investigative Ophthalmology and Visual Science*, Vol. 53, No. 1, pp. 53-61, 2012.
- [3] G. Lemaitre, M. Rastgoo, J. Massich, C.Y. Cheung, Y. Wong, E. Lamoureux, et al., "Classification of SD-OCT Volumes Using Local Binary Patterns: Experimental Validation for DME Detection," *Journal of Ophthalmology*, Vol. 6, pp. 1-16, 2016.
- [4] Y. Liu, M. Chen, H. Ishikawa, G. Wollstein, J.S. Schuman, and J.M. Rehg, "Automated Macular Pathology Diagnosis in Retinal OCT Images Using Multi-scale Spatial Pyramid with Local Binary Patterns," *Proceeding of Medical Image Computing and Computer-Assisted Intervention*, pp. 1-9, 2010.
- [5] S. Lim and D.Y. Kim, "Object Tracking Using Feature Map from Convolutional Neural Network," *Journal of Korea Multimedia Society*, Vol. 20, No. 2, pp. 126-133, 2017.
- [6] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "ImageNet: Classification with Deep Convolutional Neural Networks," *Proceeding of the*

25th International Conference on Neural Information Processing Systems, pp. 1097-1105, 2012.

[7] J.T. Lee, H. Kang, and K. Lim, "Moving Shadow Detection Using Deep Learning and Markov Random Field," *Journal of Korea Multimedia Society*, Vol. 18, No. 12, pp. 1432-1438, 2015.

[8] M.D. Abramoff, Y. Lou, and A. Erginay, "Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning," *Investigative Ophthalmology and Visual Science*, Vol. 57, No. 13, pp. 5200-5206, 2016.

[9] R. Asaoka, H. Murata, A. Iwase, and M. Araie, "Detecting Preperimetric Glaucoma with Standard Automated Perimetry Using a Deep Learning Classifier," *Ophthalmology*, Vol. 123, No. 9, pp. 1974-1980, 2016.

[10] S. Apostolopoulos, C. Ciller, S. De Zanet, S. Wolf, and R. Sznitman, "RetiNet: Automatic AMD Identification in OCT Volumetric Data," arXiv:1610.03628, 2016.

[11] T. Schlegl, S.M. Waldstein, U.M. Schmidt-Erfurth, and G. Langs, "Predicting Semantic Descriptions from Medical Images with Convolutional Neural Networks," *Information Processing in Medical Imaging*, Vol. 24, pp. 437-448, 2015.

[12] S. Cecilia, M.D. Lee, D.M. Baughman, and Y. Aaron, "Deep Learning is Effective for Classifying Normal Versus Age-related Macular Degeneration Optical Coherence Tomography Images," arXiv:1612.04891, 2016.

[13] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-scale Image Recognition," arXiv:1409.1556, 2014.

[14] S. Farsiu, S.J. Chiu, R.V. O'Connell, F.A. Folgar, E. Yuan, J.A. Izatt, et al., "Quantitative

Classification of Eyes with and without Intermediate Age-related Macular Degeneration Using Optical Coherence Tomography," *Ophthalmology*, Vol. 121, No. 1, pp. 162-172, 2014.



권 오 흠

1988년 서울대학교 컴퓨터공학과 졸업
 1991년 KAIST 전산학과 졸업 (공학석사)
 1996년 KAIST 전산학과 졸업 (공학박사)

1997년~현재 부경대학교 교수
 관심분야: 알고리즘 설계 및 분석, 의료이미지처리, 인공지능



정 유 진

1997년 서울대학교 컴퓨터공학과 학사/석사/박사 졸업
 1997년~1999년 서울대학교 컴퓨터신기술연구소 특별연구원
 1999년~2001년 부산대학교 컴퓨터 및 정보통신연구소 기금교수

2001년~현재 한국외국어대학교 컴퓨터 및 전자시스템공학부 교수
 관심분야: 모바일 보안, 무선 센서 네트워크, 의료이미지처리, 인공지능



송 하 주

1993년 서울대학교 컴퓨터공학과 졸업
 1995년 서울대학교 대학원 컴퓨터공학과 졸업(공학석사)
 2001년 서울대학교 대학원 전기 컴퓨터공학부 졸업 (공학박사)

2003년 8월 ㈜아이티포웹 부장
 2003년 9월~현재 부경대학교 교수
 관심분야: 데이터베이스 시스템, RFID/USN