

병리 인공지능 서론

송상용

성균관대학교 의과대학 삼성서울병원 병리과

Pathology has a long history of artificial intelligence (AI) as much as any other field of medicine, and has used AI algorithms continuously. However, in Korea, pathology AI is unfamiliar even to the pathologists. In this article, I will summarize the terms and definitions, the basic elements of pathology AI, and the future direction. Digital pathology is a system or environment that digitizes glass slides into binary files, observes them through a monitor or any digital devices, interprets it, analyzes it, and maintains it. Computational pathology is a comprehensive concept of diagnosis support or research system that deals with image, text and omics data. Virtual microscopy is a method or technology that allows pathologists to view and share glass slides images from whole slide scanners. Image analysis is a technique or method that processes various digital images and quantifies features. The basic elements of pathology AI are as follows: environmental factors called digital pathology and technical elements such as AI, machine learning, and deep learning. Digital pathology workflow consists of three elements; acquisition or collection of data, data processing and data storage. The basic process of image analysis consists of preprocessing of image, identification of region of interest, and feature extraction. There is enormous potential for improvement of patient care through digital pathology and/or AI, and a harmonized discussion about activation of Korean digital pathology among government, academia and industry will be mandatory for future medicine and healthcare in Korea.

Key words: Pathology; Digital Pathology; Artificial Intelligence; Machine Learning; Deep Learning

Corresponding Author: SangYong Song, M.D., Ph.D.

Department of Pathology and Translational Genomics, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Republic of Korea
Tel: +82-2-3410-2768
Fax: +82-2-3410-0025
E-mail: yoda.song@samsung.com

Received 1 Aug 2017

Revised 6 Oct 2017

Accepted 7 Nov 2017

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

서론

2016년 3월은 한국인들에게는 기억에 남을 만한 달이다. 한국 바둑계를 대표하는 이세돌 9단이 인간 대표로 구글 딥마인드의 인공지능, 즉 인공지능 대표인 알파고와 역사적인 대결을 펼쳤고, 결과는 1승 4패로 체스에 이어 바둑에서도 인공지능이 인간을 물리쳤으며[1], 1년 후 벌어진 대국에서 인공지능 알파고는 중국의 커제 9단을 3승 무패로 물리친 후 은퇴를 함으로써[2], 결국 인공지능 알파고에 승리한 최후의 인간은 한국인 이세돌로 남았기 때문이다. 구글의 알파고가 바둑으로 게임계의 최고봉을 평정하는 사이, IBM의 왓슨은 각종 인간의 활동 무대 및 산업 분야에 진출을 시도

하였고, 그 중에는 인간 활동의 최고봉이라고 해도 손색이 없을 의료계가 포함되어 있었다. 의료계를 포함시킨 이유는 아마도 첫째, 의사라는 직업이 가지는 인간 사회, 특히 지적인 면에서의 위치가 최고로 꼽아 손색이 없고, 둘째, 의료산업이 가지는 산업적 중요성 및 경제성, 셋째, 세계적인 인구 고령화에 따른 의료 수요 확대 등이 거론될 수 있다. 왓슨의 여러 가지 의료 분야 중에서도 특히 중앙학을 첫 목표로 삼았는데, 이는 암이 주는 상징성이 매우 크기 때문일 것으로 풀이될 수 있다. 미국의 암 연구 및 치료 역사는 1900년대 중반 이데올로기와 냉전 시대에 시작되어, 초기부터 외계인 또는 적과의 대립, 전쟁이라는 상징성과 형태로 프레임이 설정되었기 때문에, 암 정복 또는 암과의 전쟁이라

는 용어가 자주 사용되어 왔기 때문이다. 여하튼 의사 왓슨은 2011년 9월 웰포인트사와 파트너십을 체결한 후, 뉴욕의 메모리얼슬론캐터링 암병원에서 “수련의(resident)” 생활을 시작하였다[3]. 그 후, 왓슨은 2012년 클리블랜드 클리닉, 2013년 엠디앤더슨 암병원에서 활동을 하였고, 2016년에는 세계 진출의 일환으로 한국에까지 진출하기에 이르렀다. 인공지능은 다양한 분야의 의료계로 진출하기 시작했는데 최근 들어서는 영상 이미지를 다루는 피부과, 안과, 영상의학과 및 병리 분야의 연구가 활발해지고 있다. 이러한 인공지능의 진출에 대하여 대표적 병리학술지 중 하나인 Archives of Pathology and Laboratory Medicine에서는 2017년 5월 사설을 통해 인공지능과 병리에 대한 다양한 시각을 소개하기에 이르렀다[4-6]. 그러나, 과연 인공지능은 어느 날 갑자기 우리 옆에 나타난 것일까? 놀랍게도 병리 분야에서 인공지능에 대한 관심과 연구는 지금으로부터 30년 전으로 거슬러 올라간다.

“인간의 지능을 기계로 모방하는 과학을 인공지능이라고 하는데, 이에 대한 통일된 정의가 없기 때문에 사람들은 혼란스러워 한다. 인공지능의 정의에 대해, 혹자는 컴퓨터가 사람들이 생각하기에 지능적인 일들을 하는 것이라고 하는데 반하여, 혹자는 사람과 같은 방식으로 정보를 처리하는 컴퓨터를 인공지능이라고 한다.”

이 내용은 1985년도에 대표적인 병리학술지 중 하나인 Human Pathology에 기고된 내용에서 발췌한 것이다[7]. 당시는 인공지능의 겨울 시기였고, 저자들이 언급한 Expert system은 1965년 개발되어 연구된 대표적 알고리즘이다. 병리 분야는 의학의 다른 어느 분야 못지 않게 인공지능의 역사가 길고, 실제 알게 모르게 인공지능 알고리즘을 이용해 왔다. 이는 비단 미국의 사례뿐만 아니라 이웃 일본의 사례에서도 크게 다르지 않다. 그러나, 한국에서는 병리 인공지능이 일반인이나 타 전공분야 전문의뿐만 아니라 병리의사들에게조차 생소하다. 그래서 이번 기고를 통해 병리의사들이 알아야 할 인공지능에 대해서 지식을 공유하는 자리를 마련하고자 한다. 그러나, 병리 분야에서의 인공지능은 우리가 알고 있는 일반적인 상식에 비해서 그 역사와 범위가 길고 넓다. 따라서 본 논문에서 모든 것을 포괄하여 기술하는 것은 불가능하고, 기본이 될 수 있는 용어 및 정의, 병리 인공지능의 기본 요소 및 최근 연구 동향 등을 중심으로 정리를 해 보고자 한다.

본 론

1. 용어 및 정의

1985년 Bloom과 Weinstein의 논문[7]에서도 밝혔듯이 이 분야는 용어의 혼선이 종종 발생하는 분야이다. 따라서

용어의 정리 및 정의는 혼란을 방지하기 위해 중요하다. 이 분야의 용어는 전통적인 사전에서는 아직 찾아볼 수 없기 때문에 디지털병리학회 (Digital Pathology Association: DPA)와 최신 기술을 다루는 온라인사이트를 중심으로 정리를 하였다.

1) 디지털병리 (Digital Pathology)

DPA의 정의는 아래와 같다.

“디지털화된 유리슬라이드에서 생성된 병리 정보를 수집, 관리 및 해석할 수 있는 동적 이미지기반 환경으로써 가상현미경과 종종 같은 의미로 사용됨” (<https://digitalpathologyassociation.org/>)

한편, 프리딕셔너리 (<http://medical-dictionary.thefreedictionary.com/>)의 정의는 아래와 같다.

“조직으로부터 만든 슬라이드를 스캔하여 바이너리파일을 만든 후 컴퓨터를 이용하여 관찰하고, 저장하고, 분석하고, 보관하는 병리분석체계”

종합하여 정리하면, 디지털병리는 전통적인 병리의 유리슬라이드를 바이너리파일 형태로 디지털화 하고 이를 모니터를 통하여 관찰하고, 해석, 분석한 후 보관, 관리하는 체계 또는 환경으로 정의할 수 있다.

2) 컴퓨터병리 (Computational Pathology)

Louis 등[8]은 컴퓨터병리에 대해 아래와 같이 정의했다.

“여러 종류의 기초 자료(예: 전자병원기록(EMR), omics 데이터 및 이미지)를 통합하고, 그것으로부터 필요한 정보를 추출하고, 진단적 추론과 예측을 위해 분자, 개인, 집단 수준의 수학적 모델을 사용하고, 역동적이고 통합된 보고서와 인터페이스를 통하여 임상적으로 실행 가능한 지식을 제공하는 진단적 접근법. 이를 통하여 의사, 환자, 기사 및 기타 보건의료 관계자들이 최상의 의학적 결정을 가능하게 함.”

컴퓨터병리는 다른 어느 단어나 개념보다 최신의 개념으로써, 사전은 물론 DPA에도 정의되지 않은 분야이다. 문헌이나 검색으로 찾아 본 정의나 요약은 Louis 등의 정의에서 크게 다르지 않고, 정리하면 가상현미경이나 디지털병리가 영상 이미지 중심의 개념인데 반해, 컴퓨터병리는 영상 이미지 외에 텍스트와 오믹스 자료를 모두 포함하는 포괄적 개념의 진단 지원 체계로 볼 수 있다.

3) 가상현미경 (Virtual Microscopy)

DPA의 정의는 아래와 같다.

“네트워크를 통해 컴퓨터화면에서 볼 수 있는 현미경슬라이드로써, 현미경과 동등한 광학해상도, 시각배율, 배율변경, 초점유지가 가능하여야 함. 디지털병리와 종종 같은 의

미로 사용됨”

테크타겟 (<http://whatis.techtarget.com/>)의 정의는 아래와 같다.

“현미경이미지를 컴퓨터네트워크를 통해 보고, 전송하는 방법”

종합하여 정리하면, 가상현미경이란 전통적인 병리의 유리슬라이드를 컴퓨터화면으로 보고, 공유할 수 있는 방법, 즉 기술과 기계적 관점으로써 유리슬라이드와의 동등성이 핵심이다. 이 점을 고려하면 디지털병리의 한 요소 또는 부분집합으로 해석이 가능하다. 또한 유리슬라이드와 동등해야 하기 때문에 영상의 동적인 측면이 강조된 개념으로 볼 수 있다.

4) 이미지분석 (Image Analysis)

병리 또는 의학의 제 분야에서 이미지분석이라 함은 일반적으로 디지털이미지분석을 말하며, DPA의 정의는 아래와 같다.

“컴퓨터를 이용하여 DNA분석, 형광동소교잡반응 (Fluorescence In Situ Hybridization: FISH), 형태계측학 이미지를 처리하고 향상시켜서 특징을 탐색 (computer-assisted detection) 또는 정량 하는 방법 또는 기술”

위키피디아에서의 정의는 아래와 같다.

“이미지에서 디지털영상처리기술을 사용하여 의미 있는 정보를 추출하는 것. 컴퓨터이미지분석은 주로 컴퓨터비전 또는 기계비전, 의료영상 분야를 포함하며 패턴인식, 디지털기하학 및 신호처리를 많이 사용함”

정리하면 디지털이미지분석은 각종 디지털이미지를 처리하고 특징을 찾아내고 정량화하는 것을 의미하고, 컴퓨터공학을 이용할 경우 더 상위 개념의 구조인식을 할 수 있는 방법 또는 기술을 의미한다. 병리분야에서는 주로 정적 영상을 대상으로 연구되어 왔다.

5) 컴퓨터지원탐색 (Computer-Aided Detection/Diagnosis: CAD)

Castellino는 CAD에 대하여 아래와 같이 정의했다[9].

“거짓음성판독을 줄이기 위해 이미지의 의심스러운 특징을 찾아내서 영상의학전문의가 주목할 수 있게 해 주는 패턴인식 소프트웨어”

위키피디아에서도 비슷하게 정의하고 영상처리, 패턴인식, 인공지능이 결합된 다학제 (interdisciplinary) 기술분야라는 설명을 추가했다. 정리하면 CAD는 영상의학 분야에서 사용해 온 용어로서 병리분야의 디지털병리 또는 이미지분석에 인공지능이 결합된 형태의 기술분야로 이해할 수 있다. 이상의 용어를 정리하면 그림 1과 같다.

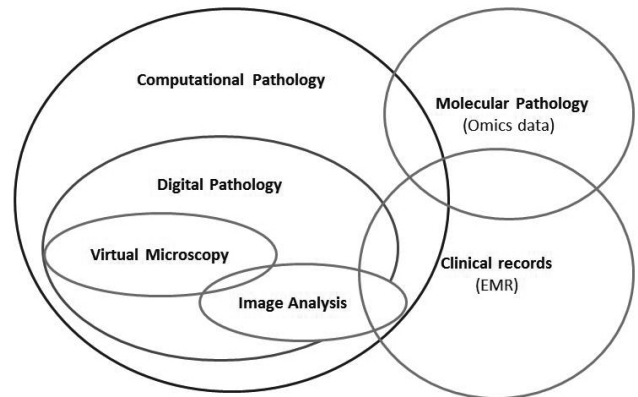


Fig. 1. Schematic illustration of new fields of pathology

2. 병리 인공지능의 요소

병리 분야에서 인공지능을 도입 또는 적용하기 위해서는 몇 가지 기본 요소들이 필요하다. 첫째는 환경적 요소로서 용어 및 정의에서 살펴본 바와 같이 디지털병리라는 환경적 요소가 필요하고, 둘째는 기술적 요소로서 컴퓨터공학/과학의 인공지능, 기계학습, 딥러닝 등의 기술적 요소가 필요하다.

1) 디지털병리

인공지능은 기본적으로 컴퓨터를 이용한 기술이자 과학 분야이기 때문에 컴퓨터가 받아들일 수 있는 자료(데이터)로 입력이 되어야 하는데, 전통적 병리는 인체로부터 채취된 조직이나 세포를 유리슬라이드에 얹어서 광학현미경으로 확대해서 관찰하는 아날로그 방식으로 업무가 이루어지기 때문에 어느 단계에서든지 디지털 이미지로 변환이 필요하다. 디지털 이미지로의 변환은 일정한 영역을 일정한 배율의 광학대 조건에서 사진을 찍는 순간촬영 (snapshot) 방식과 유리슬라이드 내용의 일부 또는 전부를 디지털화하는 전체슬라이드영상화 (whole slide imaging: WSI) 방식이 있고, 후자를 가리켜 흔히 가상현미경이라고도 부른다. 디지털병리는 정의에서도 알 수 있듯이 작업 환경으로서의 개념이 강하다. 디지털병리 환경을 위한 중요 요소는 3가지가 있는데, 첫째, 수집, 둘째, 작업, 셋째, 저장이다.

(1) 수집

디지털병리의 대상이 순간촬영이미지일 경우 수집에는 카메라가 이용될 것이고, 대상이 전체슬라이드이미지일 경우 수집에는 스캐너를 이용한다. 전자에는 유리슬라이드 정보의 극히 일부만이 수집되기 때문에 진정한 의미의 수집은 WSI로 보는 것이 타당하고, 본 논문에서는 WSI에 대해서 살펴보기로 하겠다. 첫 디지털 카메라가 1975년 코닥사의

연구원이었던 Steve Sassaon에 의해 탄생하고 2년 뒤 코니카에서 일반인을 대상으로 한 디지털카메라가 생산된 반면, 첫 디지털 스캐너의 탄생은 그로부터 20년 후인 1990년대 중후반 시작된 것으로 추정된다[10]. 실제 디지털 스캐너에 대한 아이디어 논의는 디지털 카메라의 개념과 거의 비슷한 시기에 시작되었으나 대중을 대상으로 한 카메라와 달리 스캐너의 제작 및 출시까지 20년의 격차가 있었다는 것은 기술적 어려움, 한정된 시장의 중요성을 반증한다고 볼 수 있겠다. 첫 스캐너에 대한 공식적 기록은 찾기가 어려운데, 회사들의 자료와 논문 자료가 상이하기 때문이다. 미국의 Meyer Instruments사는 1995년 첫 스캐너인 PathScan Enabler를 1995년 출시한 것으로 발표한 바 있고(unpublished data), 역시 미국의 Bacus Laboratory사는 1997년 WSI 방식의 스캐너를 출시한 것으로 알려져 있다[11]. Dunn등은 역동성 로봇 방식의 개념을 도입했다[12]. 이후, 1999년 Aperio Technologies, 2001년 헝가리 3D Histech사에서 각각 스캐너를 출시했다. 일부에서는 가상현미경이라는 이름으로도 알려졌는데 이에 대한 연구기록은 1997년 메릴랜드대학과 존스홉킨스대학병원 연구진들에 의해 발표된 바 있다[13]. 그러나 당시의 스캐너들은 느린 속도와 상대적으로 비싼 가격으로 보급에는 어려움이 있었고, 본격적인 보급은 2000년대에 이르러서 가능하게 되었다. 디지털 병리 특허 수는 1990년대 43건에서 2000년대 160건으로 400% 가까이 증가되었고, 이미지 분석 특허는 1990년대 15건에서 2000년대 39건으로 약 260% 증가한 것이 이를 뒷받침해주는 자료들이다[11]. 한국에서는 미국 연수를 다녀 온 병리과 교수들이 중심이 되어 대한병리학회 산하에 의료정보연구회를 결성하고 디지털 스캐너에 대한 관심을 가지고 있었으나 시장적인 문제로 도입이 되지 않았다. 그러나, 도입을 계속 미룰 경우 이 분야에서 국제적 수준과 격차가 커질 것을 염려한 대한병리학회에서 2008년 학회 행사인 월례집담회를 기존의 유리슬라이드 배부 방식에서 디지털 스캐너를 이용한 웹 방식으로 전격 전환한 것이 디지털병리 및 스캐너 활용의 시작이다. 현재 병리디지털스캐너를 생산하는 회사는 2000년 중반에 이미 전 세계에 걸쳐 수십 개가 넘었다[14]. 아직까지는 작업기인 뷰어의 호환성이 원활하지 않기 때문에 한 기관의 스캐너는 동일 회사 제품을 선택하는 것이 낫기 때문에 스캐너 기종 선택에 신중을 기해야 한다. 스캐너 선택 시 고려할 점은 여러 가지가 있는데, 영상의 질, 초점 층, 스캔 속도, 자동화 정도가 대표적인 고려 요소이다. 가격이 비교적 고가이기 때문에, 예산 범위에서 기관의 작업 목적에 필요한 사양을 가지고 있는 스캐너를 선택하는 것이 합리적이다. 주관적인 차이를 배제한다면 영상의 질은 광학현미경 수준에 도달했다는 것이 전문가들의 평가이고, 스캔 속도와 자동화

Table 1. Summary of whole slide scanner characteristics

Parameters	Types or characteristics
Imaging mode	Bright field (11/11) Fluorescent (3/11)
Slide capacity (slides)	4 – 400 400 (2/11) 320 (1/11) 300 (2/11) 250 (2/11) 200 (1/11) 120 (1/11) 100 (1/11)
Scan speed	20X: 30 – 90 sec 26X: 36 sec 40X: 35 – 270 sec 60X: 200 sec 33 slides/hour
Capture magnification	2X – 63X
Capture resolution	20X: 0.244 – 0.50 40X: 0.1375 – 0.25
Digital format	JPG (4/10) JPG2000 (4/10) TIFF (4/10) BigTIFF (2/10) MRXS (2/10) DICOM (1/10) VSI (1/10) RTS (1/10)
Multilayer focus	Z-stack (6/6) Extended focus (4/6)
Barcode	1D (10/10) 2D (9/10)
Special features	Automated scanning (4/10) Continuous loading (3/10) LCD touch screen (3/10) Intelligent rescan and quality scoring (1/10) Workflow acceleration (1/10) Autofluorescence reduction (1/10) Multiple remote live slide review (1/10)

정도는 해가 갈수록 발전하고 있기 때문에 향후 수년 내 스캐너가 보편화 될 수 있다는 전망들이 우세하다. 2015년 주요 스캐너 11개를 비교 검토한 자료[15]를 정리하면 표1과 같다. 유럽의 일부 병리과들은 전체 유리슬라이드를 디지털 화하여 업무에 활용하고 디지털 보관하는 사례들이 보고되고 있어서[16], 향후 병리과 운영 방식에 어떤 영향을 미칠지 주목을 받고 있다.

(2) 작업

작업의 중요한 구성 성분은 하드웨어인 모니터와 소프트웨어인 뷰어 또는 이미지분석 알고리즘/프로그램이다. 기존의 각종 논문에서 뷰어 또는 소프트웨어를 중심으로 기술한

데 반하여, 본 논문에서 작업이라는 용어를 쓴 이유는 디지털병리가 활성화될 때 가장 주목 받는 영역에 모니터가 포함될 것으로 예상되기 때문이다. 영상의학과에서는 오래 전부터 디지털 이미지를 사용하면서 디스플레이 모니터 관련한 자료와 기준들이 축적되어 왔다. 병리 이미지는 영상의학 이미지에 비해 해상도가 매우 높고, 컬러를 사용하기 때문에 영상학과 모니터 사양보다 조건들이 더 까다로울 것으로 전망이 된다. 필자의 경험으로는 일반인이나 그래픽 전문가들에게 알려진 4K 모니터 이상의 사양이 필요하다. 모니터 중에서 주목할 만한 것으로는 이른바 파워벽이라고 불리는 대형 모니터 시스템으로써 가상현실 관점의 디지털병리 분야를 구현한다[17]. 뷰어는 굳이 구분하자면 스캐너의 기본 작업 프로그램으로 이해할 수 있다. 수집기인 스캐너에서 아날로그 자료를 디지털로 변환한 파일을 모니터에 보여주는 프로그램이다. 스캐너를 만드는 회사마다 자산 스캐너에 필요한 뷰어를 만드는 게 관례였고, 그 과정에서 각 스캐너 뷰어 프로그램들은 상호 호환이 되지 않는 방향으로 발전해 왔다. 최근 주요 스캐너 회사들의 뷰어를 통합 관리하는 프로그램 개발이 시작된 것은 스캐너와 디지털병리 시장 활성화에 기여할 것으로 예상되는 고무적인 현상이다. 디지털 이미지를 모니터에 보여주는 뷰어는 관리 프로그램인 검사실정보관리시스템 (Laboratory Information Management System: LIMS)과 분석 알고리즘으로 연결되어 그 영역을 확장할 수 있다. 작업 과정을 세분화하여 살펴보면, 첫째, 이미지전처리 (image preprocessing), 둘째, 관심영역탐색 (detection of region of interest: ROI), 셋째, 특징추출 (feature extraction) 등의 단계들이 있는데, 이들은 이미지분석 목적이나 방법에 따라 다양하게 구성된다. 이런 탐색, 추출, 분석은 문제 해결이라는 과업 목적이 있기 때문에 인공지능의 관점에서 보아야 하고 후반부의 이미지분석 및 인공지능에서 다루기로 한다.

(3) 저장

고해상도로 병리 이미지를 획득할 경우, 격자 한 개 1,000 x 1,000 pixels, 슬라이드당 50 x 50개의 격자, 3 RGB로 계산하면 병리 슬라이드 한 장이 7.5 GB의 파일을 생성하게 되는 것이다. Park 등에 의하면 평균적으로 4.6 GB의 파일이 만들어진다고 했고[18], 필자의 경험에서도 3 - 5.5 GB 범위로 크게 다르지 않다. 삼성서울병원 병리과를 예로 들면, 2016년 일년 동안 헤마톡실린-에오신 (hematoxylin-eosin: HE) 염색 슬라이드 440,000장, 특수염색 80,000장이 제작되어, 만약 이를 모두 디지털 이미지로 변환한다면 약 4 PB의 디지털 이미지 파일이 생성되는데, 중요한 슬라이드에 대해 제한적으로 변환한다면 연간

축적되는 디지털 이미지 파일은 0.5-1 PB 정도로 예상된다. 그러나 유럽의 일부 병원에서는 전체 슬라이드를 일괄적으로 디지털화하여 보관하므로 정책에 따라 데이터 축적 규모는 달라질 것으로 예상된다. 디지털병리 관련 업무를 하면서 느끼는 것은 일반적으로 수집을 위한 스캐너 도입과 작업을 위한 모니터와 소프트웨어에 신경을 쓰다 보면 저장에 대해 미처 생각을 하지 못하다가 정작 작업이 순조롭게 이루어지면 그때서야 저장 대책을 새로 마련하느라 당황하게 되는 경향이 있기 때문에 전환 초기부터 중요한 요소로 점검하고 대비하여야 한다.

2) 이미지분석

병리 이미지분석에는 전통적으로 해 오던 작업으로서의 이미지분석과 최근 주목 받고 있는 딥러닝 기술을 이용한 인공지능 이미지분석이 있다. 엄밀한 의미에서 전통적인 이미지분석 역시 당시에는 인공지능에 속했었지만, 인공지능의 특성 중에는 처음 소개되었을 때는 인공지능이었다가도 보편화된 기술로 상용화될 경우 인공지능의 범주에서 독립하여 별도의 기술로 인정하는 속성이 있기에 이런 표현이 가능하다. 전술한 작업의 범위에서 소개한 바와 같이 이미지전처리, 관심영역탐색, 특징추출로 이어지는 이미지분석의 기본 과정을 살펴보기로 한다.

(1) 이미지전처리

이미지전처리는 후속작업들인 이미지분할 (image segmentation)이나 특징추출에 큰 영향을 줄 수 있다. 사실 디지털 이미지를 다루다 보면 전통적인 병리 슬라이드 제작 표준화의 중요성을 새삼 느끼게 되는데, 인체 조직의 수집 과정에서의 각종 변수들에 의한 비일관성과 인체 조직 고유의 이질성 등이 염색 슬라이드에 크고 작은 편차로 작용함을 알 수 있다. 따라서 전처리 작업은 이런 편차들을 제거 또는 감소시키는 것이 목적으로 색상정규화 (color normalization), 공간필터링 (spatial filtering), 노이즈 제거 (denoising), 대비최적화 (contrast optimization), 강도중심 (intensity centering), 히스토그램평준화 (histogram equalization) 등[19-21] 다양한 방법이 있다. 각 방법들의 상세 내용은 본 논문에서는 생략한다.

(2) 관심영역탐색

일반적인 병리 업무과정에서, 병리의사가 유리 슬라이드를 볼 때 처음 하는 일은 현미경 저배율에서 병변(관심영역)을 찾는 것인데, 이와 마찬가지로 작업기도 사실상 관심영역을 찾는 것이 실질적으로는 첫 번째 작업이다. 실제로 이 단계부터는 기계학습 또는 인공지능의 단계로 넘어가게 되기 때문에, 알고리즘 개발 때 학습이 필요한 단계이

다. 관심영역탐색은 크게 세포 수준의 탐색과 구조 수준의 탐색이 있는데 이 때 중요한 기능이 분할(segmentation)이다. 자동화된 세포 이미지 탐색 및 분할은 세포계측학에 있어서 매우 중요하고 기본적인 단계이고 각종 암 진단 및 예후 평가에 있어서도 마찬가지이다[22]. 대부분 핵 특징에 기반을 두고 있는데, 일반적으로 암 세포는 염색질 분포가 불규칙하고 핵 다형성이 있고, 림프구와 같은 종류의 염증 세포는 둥글고 작고 진하게 염색되는 핵을 가지고 있다. 핵 탐색에 사용되는 기계학습 알고리즘 방식들은 추천, LoG (Laplacian of Gaussian) 필터, 강도, 수학적 형태, H-minima transform, watershed, gradient, 색상, Markov random field (MRF), 가우시안혼합모델과 딥러닝 등이 있다[23-35]. 한편, 분할에서 많이 사용되는 모델들은 활성윤곽(active contour) 모델, 등고선(level set), 한계점(threshold), 집단화(clustering), 그래프절단(graph cuts), 분수령(watershed), 신경망 등이 있다[24,30,34-40].

(3) 특징추출

특징추출은 각종 형태학적 변수들을 세트화하는 구조적 방법이다. 일반적으로 픽셀 수준, 개체 수준 및 공간 배열 수준의 3가지 수준에서 조직의 형태를 설명한다[22]. 픽셀 수준의 특징에는 색상과 질감, 개체 수준의 특징에는 크기, 모양, 집적도, 공간배열 수준의 특징으로는 다양한 배열 모델들이 있는데, Voronoi tessellation, Delaunay triangulation, minimum spanning tree 등이 예가 될 수 있다[33,41,42]. 개체 수준의 특징과 공간배열 수준의 특징들이 비교적 형태적 개념인 크기와 모양을 기반으로 하기 때문에 이해가 비교적 쉬운 반면 픽셀 수준의 특징들은 미묘한 부분이 많다.

3. 병리 딥러닝

이미지넷 (<http://www.image-net.org/>)을 무대로 시작된 인공지능의 이미지 분석 분야의 발전은 디지털 영상 처리 및 컴퓨터 분석에 있어서 얼굴이나 사물 인식이 시작점일 것으로 생각을 하기 쉽지만, 이미지 분석의 시작은 현미경을 이용한 세포 이미지 분석이다[43]. 이런 역사적 배경답게 병리 분야에서는 컴퓨터공학의 발달과 발맞추어 많은 인공지능 알고리즘들이 사용되어 왔다. 특히 진단에 있어서 세포의 개별적 특징, 특히 핵이 강조되는 세포병리 분야에서는 수많은 연구 논문과 프로그램들이 활동을 해 왔다. 대표적인 사례로 자궁경부암세포검사인 액상세포검사에서 기계학습 알고리즘을 기반으로 하는 자동 선별검사가 보완적으로 사용할 수 있는 것으로 알려져 있고[44,45], 연구 논문은 수백 편에 이르고 있다. 유방암조

직병리는 인공지능과 기계학습의 현미경 이미지 연구의 메카라고 해도 과언이 아닐 만큼 다양하고 오랜 연구의 역사를 가지고 있다. 그 이유는 바로 환자의 진단, 치료, 예후에 있어서 조직병리 소견이 절대적으로 중요한데, 이 소견이 반정량적(semi-quantitative)이기 때문에 병리의사의 노력이 많이 필요하고, 관찰자간 및 관찰자내 편차가 존재하기 때문에 보완 방법의 필요성이 제기되어 왔기 때문이다. 유방암의 조직병리 중 기계학습 분야에서 관심을 가져온 분야는 첫째, 면역조직화학염색판독[46,47], 둘째, 유사분열[48-52], 셋째, 종양침윤립프구[53-55]가 대표적이다. 물론 이외에도 종양의 종류나 형태에 관한 다른 측면의 연구들도 시행되어 왔다. 유방암의 조직병리 등급체계인 Nottingham grading system은 세관형성, 핵 다형성, 유사분열수의 세 가지 변수로 구성되는데, 그 중에서도 특히 유사분열은 오랫동안 기계학습 전문가들의 관심과 연구대상이었다. 수작업으로 유사분열을 세는 것은 주관적이기 때문에 관찰자내, 관찰자간 편차도 심하고, 현미경 고배율 검경이 필요하여 피로도가 높으며, 단순한 작업의 특징인 지루한 작업으로 받아들여져 왔다. 자동화 유사분열 계측으로 경연을 벌이는 International Conference on Pattern Recognition (ICPR)과 Medical Image Computing and Computer-Assisted Intervention (MICCAI)에서 여러 기계학습 중에서 심층신경망(딥러닝)이 가장 우수한 성적을 보였고, 그 정확도가 숙련된 병리의사와 유사한 수준으로 나타났다[49,50, 56,57]. 현재 인공지능을 이용한 병리 관련 연구는 크게 이미지 연구와 보고서 연구로 나눌 수 있다. 보고서 연구는 한마디로 설명하자면 “왓슨병리”라는 표현이 적당할 듯 하다. 이미지 연구는 컴퓨터하드웨어의 고성능화, 딥러닝 기술의 도입 등으로 기대감이 높아졌으나, 연구를 위한 데이터세트의 부족으로 장애가 있는 게 현실이다. 더구나 한국의 경우는 디지털병리 기반이 매우 취약하여 국제 경쟁력에 허점을 보이고 있기 때문에 정부의 정책적, 제도적 뒷받침이 필요한 시점이다. 영국은 오랜 준비 끝에 2014년 “국가병리프로그램 디지털최우선: 병리혁신을 통한 임상변화(National Pathology Programme Digital First: Clinical Transformation through Pathology Innovation)”라는 국가 차원의 디지털병리 사업을 선언했다. 영국 정부가 환자들에게 제공하는 진료의 질은 디지털병리를 통해서만 가능하다고 판단했기 때문이다. 병리와 인공지능에 대한 기고를 의뢰 받으며 인공지능 중심으로 풀어야 할지, 디지털병리 중심으로 풀어야 할지 오랜 고민 끝에 디지털병리에 무게중심을 둔 이유도 여기에 있다. 좋은 유람선이 있어도 강에 물이 마르면 의미가 없듯이 병리의 디지털화 없이는 정밀의학도 인공지능의 발전도 국제경쟁력을 가질 수 없다고 필자는 판단한다. 자원 없이 인재로 경쟁

하는 한국이 중요한 국제경쟁력 한 가지를 잃지 않기를 바란다.

결 론

지금은 인공지능의 시대이다. 의료 분야는 고령화와 각종 질병의 증가로 나날이 증가하는 의료 수요와 함께 환자에게 치명적일 수 있는 오류 감소라는 명제, 그리고 의료 인력의 효율적 활용을 위하여 인공지능의 수요가 매우 높다. 미국과 일본은 30년이 넘는 인공지능 강국이고, 영국은 국가 차원의 디지털병리를 선언하여 실행에 옮기고 있으며, 중국은 다른 어느 나라와도 비교할 수 없는 속도의 경쟁을 시도하고 있다. 유럽의 병리과들은 유리슬라이드 대신 디지털슬라이드를 업무에 적용하는 파격적 변신을 시도하고 있는 이 시점에 한국에는 스캐너 한 대 없는 병리과가 대부분이다. 2008년도 대한병리학회 월례집담회를 디지털 슬라이드로 전환한 이후 한국에는 전문가들조차 디지털병리와 관련된 아무 움직임이 없다. 인공지능 병리를 통한 진단 지원 및 질병 예측은 미래 의료로 나아가는 중요한 발판이 될 것이라는 것을 다른 나라의 사례를 통하여 지금이라도 자각하고, 정부와 전문가단체와 산업계가 중지를 모아 더 나은 미래 의료를 위해 디지털병리와 병리인공지능을 활성화 해야 할 시점이다.

REFERENCES

1. Borowiec S. AlphaGo seals 4-1 victory over Go grandmaster Lee Sedol. The Guardian 15 March, 2016.
2. Byford S. AlphaGo retires from competitive Go after defeating world number one 3-0. The Verge 27 May, 2017
3. Fitzgerald J. IBM, NYC hospital training Watson supercomputer in cancer. Phys Org 22 March, 2012
4. Granter SR, Beck AH, Papke DJ Jr. AlphaGo, Deep Learning, and the Future of the Human Microscopist. Arch Pathol Lab Med 2017;141:619-21.
5. Sharma G, Carter A. Artificial Intelligence and the Pathologist: Future Frenemies? Arch Pathol Lab Med 2017;141:622-3.
6. Granter SR, Beck AH, Papke DJ Jr. Straw Men, Deep Learning, and the Future of the Human Microscopist: Response to "Artificial Intelligence and the Pathologist: Future Frenemies?" Arch Pathol Lab Med 2017;141:624.
7. Bloom KJ and Weinstein RS. Expert systems: robot physicians of the future? Hum Pathol 1985;16:1082-4.
8. Louis DN, Gerber GK, Baron JM, Bry L, Dighe AS, Getz G, et al. Computational Pathology. An Emerging Definition. Arch Pathol Lab Med 2014;138:1133-8.
9. Castellino RA. Computer aided detection (CAD): an overview. Cancer

- Imaging 2005;5:17-9.
10. Song SY. Artificial Intelligence in Pathology. The 25th federation meeting of Korean basic medical scientists 2017. S33, 2017.
11. Cucoranu IC, Parwani AV, Vepa S, Weinstein RS, Pantanowitz L. Digital pathology: A systematic evaluation of the patent landscape. J Pathol Inform 2014;1:16.
12. Dunn BE, Almagro UA, Choi H, Sheth NK, Arnold JS, Recla DL, et al. Dynamic-robotic telepathology: Department of Veterans Affairs feasibility study. Hum Pathol 1997;28:8-12.
13. Ferreirat R, Moon B, Humphriest J, Sussman A, Saltz J, Miller R, et al. The Virtual Microscope. Proc AMIA Annu Fall Symp 1997:449-53.
14. Rojo MG, García GB, Mateos CP, García JG, Vicente MC. Critical comparison of 31 commercially available digital slide systems in pathology. Int J Surg Pathol 2006;14:285-305.
15. Farahani N, Parwani AV, Pantanowitz L. Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. Pathol Lab Med Int 2015;7:23-33.
16. Baidoshvili A. Making the Move to 100 Percent Digital. The Pathologist 2015:502-3.
17. Randell R, Ruddle RA, Mello-Thoms C, Thomas RG, Quirke P, Treanor D. Virtual reality microscope versus conventional microscope regarding time to diagnosis: An experimental study. Histopathology 2013;62:351-8.
18. Park S, Pantanowitz L, Parwani AV. Digital imaging in pathology. Clin Lab Med. 2012;32:557-84.
19. Bhattacharyya S. A brief survey of color image preprocessing and segmentation techniques. J Pattern Recognit Res 2011;6:120-9.
20. Khan AM, Rajpoot N, Treanor D, Magee D. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. IEEE Trans Biomed Eng 2014;61:1729-38.
21. He L, Long LR, Antani S, Thoma GR. Histology image analysis for carcinoma detection and grading. Comput Methods Programs Biomed 2012;107:538-56.
22. Chen JM, Li Y, Xu J, Gong L, Wang LW, Liu WL, Liu J. Computer-aided prognosis on breast cancer with hematoxylin and eosin histopathology images: A review. Tumour Biol 2017;39:1-12.
23. Parvin B, Yang Q, Han J, Chang H, Rydberg B, Barcellos-Hoff MH. Iterative voting for inference of structural saliency and characterization of subcellular events. IEEE Trans Image Process 2007;16:615-23.
24. Al-Kofahi Y, Lassoued W, Lee W, Roysam B. Improved automatic detection and segmentation of cell nuclei in histopathology images. IEEE Trans Biomed Eng 2010;57:841-52.
25. Arteta C, Lempitsky V, Noble JA, Zisserman A. Learning to detect cells using non-overlapping extremal regions. Med Image Comput Comput Assist Interv 2012;7510:348-56.
26. Yan C, Li A, Zhang B, Ding W, Luo Q, Gong H. Automated and accurate

- detection of soma location and surface morphology in large-scale 3D neuron images. *PLoS One* 2013;8:e62579.
27. Esmailsabzali H, Sakaki K, Dechev N, Burke RD, Park EJ. Machine vision-based localization of nucleic and cytoplasmic injection sites on low-contrast adherent cells. *Med Biol Eng Comput* 2012;50:11–21.
 28. Jung C and Kim C. Segmenting clustered nuclei using H-minima transform-based marker extraction and contour parameterization. *IEEE Trans Biomed Eng* 2010;57:2600–04.
 29. Li F, Zhou X, Ma J, Wong ST. Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis. *IEEE Trans Med Imaging* 2010;29:96–105.
 30. Ali S and Madabhushi A. An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery. *IEEE Trans Med Imaging* 2012;31:1448–60.
 31. Cheng JR and Rajapakse J. Segmentation of clustered nuclei with shape markers and marking function. *IEEE Trans Biomed Eng* 2009;56:741–8.
 32. Basavanthally AN, Ganesan S, Agner S, Monaco JP, Feldman MD, Tomaszewski JE, et al. Computerized image-based detection and grading of lymphocytic infiltration in her2+ breast cancer histopathology. *IEEE Trans Biomed Eng* 2010;57:642–53.
 33. Fatakawala H, Xu J, Basavanthally A, Bhanot G, Ganesan S, Feldman M, et al. Expectation-maximization-driven geodesic active contour with overlap resolution (EMAGACOR): application to lymphocyte segmentation on breast cancer histopathology. *IEEE Trans Biomed Eng* 2010;57:1676–89.
 34. Vink JP, Van Leeuwen MB, Van Deurzen CH, De Haan G. Efficient nucleus detector in histopathology images. *J Microsc* 2013;249:124–35.
 35. Xu J, Xiang L, Liu Q, Gilmore H, Wu J, Tang J, Madabhushi A. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. *IEEE Trans Med Imaging* 2016;35:119–30.
 36. Hoque A, Lippman SM, Boiko IV, Atkinson EN, Sneige N, Sahin A, et al. Quantitative nuclear morphometry by image analysis for prediction of recurrence of ductal carcinoma in situ of the breast. *Cancer Epidemiol Biomarkers Prev* 2001;10:249–59.
 37. Qi X, Xing F, Foran DJ, Yang L. Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. *IEEE Trans Biomed Eng* 2012;59:754–65.
 38. Latson L, Sebek B, Powell KA. Automated cell nuclear segmentation in color images of hematoxylin and eosin-stained breast biopsy. *Anal Quant Cytol Histol* 2003;25:321–31.
 39. Veta M, Kornegoor R, Huisman A, Verschuur-Maes AH, Viergever MA, Pluim JP, van Diest PJ. Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. *Mod Pathol* 2012;25:1559–65.
 40. Mouelhi A, Sayadi M, Fnaiech F. A supervised segmentation scheme based on multilayer neural network and color active contour model for breast cancer nuclei detection. In: 2013 international conference on electrical engineering and software applications (ICEESA), Hammamet, Tunisia, 21–23 March 2013;53:1–6. New York: IEEE.
 41. Naik S, Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski J. Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. In: 5th IEEE international symposium on biomedical imaging: from nano to macro, Paris, France, 14–17 May 2008, pp. 284–287. New York: IEEE.
 42. Basavanthally A, Ganesan S, Feldman M, Shih N, Mies C, Tomaszewski J, et al. Multifield-of-view framework for distinguishing tumor grade in ER+ breast cancer from entire histopathology slides. *IEEE Trans Biomed Eng* 2013; 60(8): 2089–2099.
 43. Mendelsohn ML, Kolman WA, Perry B, Prewitt JM. Morphological analysis of cells and chromosomes by digital computer. *Methods Inf Med* 1965;4:163–7.
 44. Wilbur DC, Prey MU, Miller WM, Pawlick GF, Colgan TJ. The AutoPap system for primary screening in cervical cytology. Comparing the results of a prospective, intended-use study with routine manual practice. *Acta Cytol* 1998;42:214–20.
 45. Andrea E. Dawson. Can We Change the Way We Screen?: The ThinPrep Imaging System. *Cancer Cytopathol* 2004;102:340–4.
 46. Mungle T, Tewary S, DAS DK, Arun I, Basak B, Agarwal S, et al. MRF-ANN: a machine learning approach for automated ER scoring of breast cancer immunohistochemical images. *J Microsc* 2017;267:117–29.
 47. Tuominen VJ, Ruotoistenmäki S, Viitanen A, Jumppanen M, Isola J. ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67. *Breast Cancer Res* 2010;12:R56.
 48. Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci S, Navab N. AggNet: Deep Learning From Crowds for Mitosis Detection in Breast Cancer Histology Images. *IEEE Trans Med Imaging* 2016;35:1313–21.
 49. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks. *Med Image Comput Comput Assist Interv* 2013;16:411–8.
 50. Veta M, van Diest PJ, Willems SM, Wang H, Madabhushi A, Cruz-Roa A, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med Image Anal* 2015;20:237–48.
 51. Wang H, Cruz-Roa A, Basavanthally A, Gilmore H, Shih N, Feldman M, et al. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *J Med Imaging (Bellingham)* 2014;1:034003.
 52. Malon CD, Cosatto E. Classification of mitotic figures with convolutional neural networks and seeded blob features. *J Pathol Inform.* 2013;4:9.
 53. Ali HR, Dariush A, Provenzano E, Bardwell H, Abraham JE, Iddawela

- M, et al. Computational pathology of pre-treatment biopsies identifies lymphocyte density as a predictor of response to neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res* 2016;18:21.
54. Turkki R, Linder N, Kovanen PE, Pellinen T, Lundin J. Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. *J Pathol Inform* 2016;7:38.
55. Yuan Y. Modelling the spatial heterogeneity and molecular correlates of lymphocytic infiltration in triple-negative breast cancer. *J R Soc Interface* 2015;12:20141153.
56. Roux L, Racoceanu D, Loménie N, Kulikova M, Irshad H, Klossa J, et al. Mitosis detection in breast cancer histological images an ICPR 2012 contest. *J Pathol Inform* 2013;4:8.
57. Giusti A, Cacciay C, Ciresan DC, Schmidhuber J, Gambardella LM. A comparison of algorithms and humans for mitosis detection. In: 2014 IEEE 11th international symposium on biomed imaging (ISBI), Beijing, China, 29 April–2 May 2014, pp. 1360–1363. New York: IEEE.