



상부위장관 암 진단을 위한 인공지능 분석, 어디까지 발전하였는가?

이정수¹, 김남국², 김도훈³인제대학교 의과대학 일산백병원 소화기내과¹, 울산대학교 의과대학 서울아산병원 융합의학과², 소화기내과³

To What Degree Has Artificial Intelligence Developed for Diagnosis of Upper Gastrointestinal Cancer?

Jung Su Lee¹, Namkug Kim², Do Hoon Kim³Division of Gastroenterology, Department of Internal Medicine, Ilsan Paik Hospital, Inje University College of Medicine¹, Goyang, Department of Convergence Medicine², Division of Gastroenterology, Department of Internal Medicine³, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea**Article:** Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study (*Lancet Oncol* 2019;20:1645-1654)

요약: 식도암과 위암을 포함한 상부위장관 암은 전 세계적으로 가장 흔한 암이며 암 관련 사망률 또한 높다. 진행성 암의 경우 예후가 매우 불량하지만 조기에 진단된 경우는 5년 생존율이 90%에 육박한다. 조기 위장관 암을 진단하기 위하여 협대역 영상 기술 등 관련 노력이 지속되고 있으나 종종 진단에 어려움을 겪고 있는 실정이다. 최근 많은 의료 분야에서 인공지능을 기반으로 한 영상 분석이 사람의 눈으로 확인하기 어려운 구조나 특징까지 분석할 수 있는 것으로 보고하고 있다. 상부위장관 암을 조기에 진단하기 위하여 이번 연구는 내시경 영상에 대하여 인공지능을 기반으로 한 진단 방법을 개발하고 그 효과를 확인하기 위하여 시행되었다.

중국 내 6개 병원의 총 84,424명에서 촬영된 1,036,496장의 백색광 내시경 영상을 연구에 활용하여 인공지능 기반 진단 프로그램을 개발하고 검증하였다. 후향적으로 18,765명에서 촬영된 157,207장의 백색광 내시경 영상을 수집한 후 이를 Google에서 배포한 DeepLab's V3+ 기반의 인공지능에 교육시켜 Gastrointestinal Artificial Intelligence Diagnostic System (GRAIDS)을 개발하였다. 그리고 GRAIDS는 2018년 7월 21일에 온라인으로 공개하였으며 Sun Yat-sen 대학병원의 암센터를 포함한 다른

5개의 병원에서 사용 중인 내시경 기계에 GRAIDS를 추가 장착하여 전향적으로 모집된 환자 65,659명을 대상으로 내시경을 통하여 실시간으로 촬영한 879,289장의 백색광 내시경 영상을 분석하여 식도암과 위암 진단에 대한 효과를 확인하였다. 인공지능의 성능을 비교하기 위하여 내시경 시행 의사의 경력별로 10년 이상의 전문가 집단, 5년 이상의 준전문가 집단, 2년 이내의 수련생 집단으로 구분하여 네 집단 간의 진단 정확도, 민감도, 특이도, 양성 예측률, 음성 예측률의 차이를 확인하였다. 진단 정확도에 있어서 GRAIDS가 전문가 집단에 비하여 유의하게 다소 낮았으나(0.928 [95% CI 0.919~0.937] vs. 0.967 [95% CI 0.961~0.973]; $P < 0.0001$), 민감도에서는 유의한 차이가 없었다(0.942 [95% CI 0.924~0.957] vs. 0.945 [95% CI 0.927~0.959]; $P = 0.692$). 준전문가 집단(0.858 [95% CI 0.832~0.880], $P < 0.0001$) 및 수련생 집단(0.722 [95% CI 0.691~0.752], $P < 0.0001$)과 비교하였을 때는 GRAIDS의 민감도가 우수한 것으로 확인하였다. 양성 예측률에서 GRAIDS는 0.814 (95% CI 0.788~0.838), 전문가 집단은 0.932 (95% CI 0.913~0.948), 준전문가 집단은 0.974 (95% CI 0.960~0.984), 수련생 집단은 0.824 (95% CI 0.795~0.850)로 확인되었으며, GRAIDS가 전문가 집단과 준전문가 집단에 비해서는 낮았으나 수련생 집단과는 비슷한 수준으로 확인되었다. GRAIDS의 특이도는 0.923 (95% CI 0.912~0.933), 음성 예측률은 0.978 (95% CI 0.971~0.984)로 다른 집단들과 비슷한 수준으로 확인되었다. 특히, 전문가 집단과 GRAIDS를 조합할 경우에 민감도는

Received: February 10, 2020 Revised: February 29, 2020 Accepted: March 5, 2020

Corresponding author: Do Hoon Kim

Division of Gastroenterology, Department of Internal Medicine, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea

Tel: +82-2-3010-3193, Fax: +82-2-476-0824, E-mail: dohoon.md@gmail.com

Copyright © 2020 Korean College of *Helicobacter* and Upper Gastrointestinal Research© The Korean Journal of *Helicobacter* and Upper Gastrointestinal Research is an Open-Access Journal. All articles are distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

GRAIDS 혹은 전문가 집단에 비하여 유의하게 높아지는 것으로 확인되었다(0.984 [95% CI 0.973~0.991], $P<0.0001$). 준전문가 집단(0.978 [95% CI 0.966~0.987], $P<0.0001$)과 수련생 집단(0.964 [95% CI 0.949~0.975], $P<0.0001$)에 GRAIDS를 조합한 경우에서도 민감도의 의미 있는 향상이 확인되었다.

결론적으로 인공지능 기반으로 개발된 상부위장관 암 진단 프로그램인 GRAIDS는 높은 진단 정확도를 보이고 있으며, 민감도는 전문가 집단과 비슷한 수준임을 확인하였다. 이를 지역사회 병원에서 활용하게 될 경우 상부위장관 암 진단 능력의 의미 있는 향상을 기대해볼 수 있을 것이다.

해설: 2012년 딥러닝(deep learning)이란 방법론을 통하여 세계 영상 인식 대회(ImageNet Challenge)에서 토론토 대학교의 연구팀이 우승하면서 전 세계적으로 인공지능 연구는 새로운 국면을 맞이하게 되었다. 이후 인공지능 관련 기술들은 하루가 다르게 발전하였고, 의료에서는 영상의학적인 진단뿐만 아니라 의무기록을 기반으로 한 예후 예측 모델 등에서 연구가 활발히 이루어졌고 그 성능을 보고한 바 있다.^{1,2} 최근 소화기내과의 진단 내시경 분야에서도 인공지능을 기반으로 하는 다양한 연구 결과들이 발표되고 있다. 특히, 실시간으로 대장 용종의 유무를 인공지능 기반 기술을 통하여 진단하여 선종 발견율이 유의하게 향상되었으며(29.1% vs. 20.3%, $P<0.001$), 나이가 해당 용종이 비종양성인지 종양성인지를 예측하는 수준까지 보고되었다.^{3,4} 이러한 연구 개발을 토대로 2019년 3월 4일 올림푸사를 통하여 EndoBRAIN이라 불리는 인공지능 기반 진단 내시경 기술이 최초로 상업화에 성공하였다.⁵

상부위장관 질환에 대한 연구도 최근에 활발히 이루어지고 있다. Zhu 등⁶은 인공지능을 기반으로 한 백색광 내시경 영상 분석을 통하여 위암의 침윤 예측이 내시경 시행 의사보다 인공지능이 우수한 것으로 보고하였고, Yoon 등⁷은 SELVAS AI사와 협업을 통하여 백색광 내시경 영상으로 인공지능 기반 조기 위암 진단 및 침윤 예측의 우수한 성능을 보고한 바 있다. 또한, Cho 등⁸은 위암을 진단함에 있어 중요한 감별진단인 양성 위궤양 등을 포함한 비종양성 병변, 저도 이형성증, 고도 이형성증과 조기 위암, 진행성 위암을 인공지능을 기반으로 하여 감별진단이 가능할 수 있음을 보고하였다. 다만, Lee 등⁹은 정상과 양성 위궤양, 정상과 위암을 감별진단하는 데 인공지능 기반 분석의 정확도가 각각 96%, 92%임을 확인하였으나 임상적 난제인 양성 위궤양과 위암의 감별진단 정확도는 77%로 인공지능으로 감별진단하기 쉽지 않았음을 보고하였다. 위암뿐만 아니라 식도암 진단에 대해서도 연구가 보고되고 있다. Ohmori 등¹⁰과 Nakagawa 등¹¹이 표재성 식도암에 대하여 백색광 내시경 영상에 확대내시경 영상과 협대역 영상을 포함한 인공지능 기반 분

석이 전문가 수준의 내시경 시행 의사와 비교하였을 때 표재성 식도암 진단 및 침윤 예측에 대하여 비슷한 정도의 성능을 보이고 있음을 보고하였다.

이러한 연구를 통하여 이루어지고 있는 인공지능 기술의 발전은 내시경 분야에서 조기 상부위장관 암의 진단에 큰 도움이 될 것으로 기대한다. 나아가 조기 진단 이후 암의 침윤을 예측하여 치료 내시경의 적응증을 확인하고 조기 상부위장관 암 환자에게 치료 내시경 시술의 기회를 제공할 수 있을 것이며, 상부위장관 암의 전체적인 예후 향상에도 도움이 될 것으로 기대할 수 있다. 또한, 인공지능 기반 진단 내시경 기술의 발전은 내시경 시행 의사의 과중한 업무 피로도를 줄여줄 것으로 기대된다. 특히, 이번에 소개한 연구는 백색광 내시경 영상만을 이용하여 분석한 것으로, 협대역 영상 기술 등의 특수 영상이 보급되지 않은 곳이나 비전문가가 내시경을 시행할 때 더욱 유용할 것으로 생각한다. 다만, 이번 연구는 다음과 같은 제한점들이 있다. 첫째, 중국인들의 자료로만 분석되었기에 여러 나라에 적용하는 데 제한이 있을 수 있다. 예를 들어, 식도 종양에 있어 서양은 바렛식도에 의한 종양이 흔하고 동양은 편평세포 종양이 흔하다. 그럼에도 불구하고 중국 내 6개 병원의 대규모 자료를 활용하여 해당 결과를 보고한 것은 상당히 의미 있는 연구라 할 수 있다. 둘째, GRAIDS의 양성 예측률이 0.814 (95% CI 0.788~0.838)로 전문가 집단(0.932 [95% CI 0.913~0.948])에 비하여 낮고 수련생 집단(0.824 [95% CI 0.795~0.850])과 비슷한 수준으로 확인되었다. 이러한 문제는 주로 GRAIDS가 정상 구조물(유문부, 위각부, 다량의 점액질, 연동운동 중인 위벽 등)을 병변으로 오인한 경우와 상부위장관 암의 유병률이 낮은 자료(3.8~9.5%)로 검증을 진행한 경우 나타났다. 내시경 시행 당시 정상 구조물들은 내시경 시행 의사가 쉽게 구별하여 위양성임을 확인할 수 있을 것으로 생각하며, GRAIDS에 의하여 상대적으로 위양성이 증가하더라도 이로 인하여 암을 진단하지 못하는 경우를 줄여줄 것으로 기대한다.

전 세계적으로 다양한 분야에서 인공지능을 접목하여 여러 가지 편의성을 제공하며 나날이 발전하고 있다. 의료 분야 또한 다양한 의무기록 및 영상의학적 정보를 토대로 디지털 정보의 과학화가 이루어지고 있으며, 향후 5~10년 후에는 새로운 국면을 맞이할 수도 있을 것으로 생각한다. 따라서, 우리는 앞으로 보고될 인공지능 기반 기술을 활용한 후속 임상 연구들에 대하여 적극적인 관심을 가지고 신중히 접근하여 그 결과를 확인하여야 할 것이다.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article

was reported.

ORCID

Jung Su Lee  <https://orcid.org/0000-0001-5644-0150>
Namkug Kim  <https://orcid.org/0000-0002-3438-2217>
Do Hoon Kim  <https://orcid.org/0000-0002-4250-4683>

REFERENCES

1. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. *Radiographics* 2017;37:505-515.
2. Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine learning-based model for prediction of outcomes in acute stroke. *Stroke* 2019;50:1263-1265.
3. Wang P, Berzin TM, Glissen Brown JR, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019;68:1813-1819.
4. Song EM, Park B, Ha CA, et al. Endoscopic diagnosis and treatment planning for colorectal polyps using a deep-learning model. *Sci Rep* 2020;10:30.
5. Yamada T. Development of an open platform for AI-assisted diagnosis in endoscopic examinations [Internet]. Tokyo: OLYMPUS [2019 Mar 13; cited 2020 Jan 31]. Available from: https://www.olympus-global.com/ir/data/pdf/ict-ai_platform_2019e_03.pdf.
6. Zhu Y, Wang QC, Xu MD, et al. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. *Gastrointest Endosc* 2019;89:806-815.e1.
7. Yoon HJ, Kim S, Kim JH, et al. A Lesion-based convolutional neural network improves endoscopic detection and depth prediction of early gastric cancer. *J Clin Med* 2019;8:1310.
8. Cho BJ, Bang CS, Park SW, et al. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* 2019;51:1121-1129.
9. Lee JH, Kim YJ, Kim YW, et al. Spotting malignancies from gastric endoscopic images using deep learning. *Surg Endosc* 2019;33:3790-3797.
10. Ohmori M, Ishihara R, Aoyama K, et al. Endoscopic detection and differentiation of esophageal lesions using a deep neural network. *Gastrointest Endosc* 2020;91:301-309.e1.
11. Nakagawa K, Ishihara R, Aoyama K, et al. Classification for invasion depth of esophageal squamous cell carcinoma using a deep neural network compared with experienced endoscopists. *Gastrointest Endosc* 2019;90:407-414.