



다양한 임상 시나리오에서의 이미지 기반 예후 예측

성명 황의진

소속 서울대학교병원

20 제4회
25 COMMUNITY
CONFERENCE

TANGO

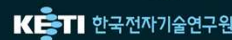
Target Adaptive No-code neural network
Generation and Operation framework

주 관 ETRI (TANGO)

주 최 과학기술정보통신부 IITP 정보통신기획평가원

문의 parkjb@etri.re.kr / 042-860-5565

후 원



iRAIL 소개

iRAIL (innovative Radiology Artificial Intelligence Lab)

- 서울대병원의 임상 의사 및 서울대 공대의 AI 엔지니어로 이루어진 의료영상인공지능 연구실
- 흉부영상의학에서 시작해 의료영상 전 범위에 대한 연구로 확장
- 세계 최초로 딥러닝 기반의 흉부 X-ray 진단 연구를 수행 및 발표



- Director: 박창민 교수님
- TANGO 활용연구 PI: 황의진 교수님



iRAIL 소개

200+ SCI papers, 40+ patents

Radiology

Performance of a Deep Learning Algorithm Compared with Radiologic Interpretation for Lung Cancer Detection on Chest Radiographs in a Health Screening Population

Jung Hyuk Lee, MD* • Hye Young Sun, MD* • Sungyeon Park, PhD • Hyungjin Kim, MD, PhD • Eui Jin Hwang, MD • Jin Mo Goo, MD, PhD • Chang Min Park, MD, PhD

From the Department of Radiology and Institute of Radiation Medicine, Seoul National University College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea (J.H.L., H.Y.S., E.J.H., J.M.G., C.M.P.); Department of Radiology, Healthcare System Gangnam Center, Seoul National University Hospital, Seoul, Korea (H.Y.S.); and Lunit Inc, Seoul, Korea (S.P.). Received March 31, 2020; revision received July 20; accepted July 30. Address correspondence to C.M.P. (e-mail: cnpark.morphus@gmail.com).

This study was supported by the Seoul National University Hospital Research Fund (grant no. 03-2019-0190).

*J.H.L. and H.Y.S. contributed equally to this work.

Conflicts of interest are listed at the end of this article.

See also the editorial by Armento in this issue.

Radiology 2020; 297:687–696 • <https://doi.org/10.1148/radiol.2020201240> • Content codes: CH DI

Background: The performance of a deep learning algorithm for lung cancer detection on chest radiographs in a health screening population is unknown.

Purpose: To validate a commercially available deep learning algorithm for lung cancer detection on chest radiographs in a health screening population.

Materials and Methods: Out-of-sample testing of a deep learning algorithm was retrospectively performed using chest radiographs from individuals undergoing a comprehensive medical check-up between July 2008 and December 2008 (validation test). To evaluate the algorithm performance for visible lung cancer detection, the area under the receiver operating characteristic curve (AUC) and diagnostic measures, including sensitivity and false-positive rate (FPR), were calculated. The algorithm performance was compared with that of radiologists using the McNemar test and the Moskowitz method. Additionally, the deep learning algorithm was applied to a screening cohort undergoing chest radiography between January 2008 and December 2012, and its performances were calculated.

Results: In a validation test comprising 10 285 radiographs from 10 202 individuals (mean age, 54 years \pm 11 [standard deviation]; 5857 men) with 10 radiographs of visible lung cancers, the algorithm's AUC was 0.99 (95% confidence interval, 0.97, 1), and it showed comparable sensitivity (90% [nine of 10 radiographs]) to that of the radiologists (60% [six of 10 radiographs]; $P = .25$) with a higher FPR (3.1% [319 of 10 275 radiographs] vs 0.3% [26 of 10 275 radiographs]; $P < .001$). In the screening cohort of 100 525 chest radiographs from 50 070 individuals (mean age, 53 years \pm 11; 28 090 men) with 47 radiographs of visible lung cancers, the algorithm's AUC was 0.97 (95% confidence interval, 0.95, 0.99), and its sensitivity and FPR were 83% (39 of 47 radiographs) and 3% (2997 of 100 478 radiographs), respectively.

Conclusion: A deep learning algorithm detected lung cancers on chest radiographs with a performance comparable to that of radiologists, which will be helpful for radiologists in healthy populations with a low prevalence of lung cancer.

©RSNA, 2020

Online supplemental material is available for this article.

ORIGINAL RESEARCH • THORACIC IMAGING

JAMA Network Open

Development and Validation of a Deep Learning–Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs

Eui Jin Hwang, MD, Sungyeon Park, MS, Kwang Nam Jin, MD, Jung In Kim, MD, Se Young Choi, MD, Jong Hyuk Lee, MD, Jin Mo Goo, MD, PhD, Jaehong Aum, PhD, Jae-Joon Yim, MD, Julien G. Cohen, MD, Gilbert R. Ferretti, MD, Chang Min Park, MD, PhD, for the DLAD Development and Evaluation Group

Abstract

IMPORTANCE: Interpretation of chest radiographs is a challenging task prone to errors, requiring expert readers. An automated system that can accurately classify chest radiographs may help streamline the clinical workflow.

OBJECTIVES: To develop a deep learning–based algorithm that can classify normal and abnormal results from chest radiographs with major thoracic diseases including pulmonary malignant neoplasm, active tuberculosis, pneumonia, and pneumothorax and to validate the algorithm's performance using independent data sets.

DESIGN, SETTING, AND PARTICIPANTS: This diagnostic study developed a deep learning–based algorithm using single-center data collected between November 1, 2016, and January 31, 2017. The algorithm was externally validated with multicenter data collected between May 1 and July 31, 2018. A total of 54 221 chest radiographs with normal findings from 47 917 individuals (21 556 men and 26 361 women; mean [SD] age, 51 [16] years) and 35 613 chest radiographs with abnormal findings from 14 102 individuals (8373 men and 5729 women; mean [SD] age, 62 [15] years) were used to develop the algorithm. A total of 486 chest radiographs with normal results and 529 with abnormal results (1 from each participant; 628 men and 387 women; mean [SD] age, 53 [18] years) from 5 institutions were used for external validation. Fifteen physicians, including nonradiology physicians, board-certified radiologists, and thoracic radiologists, participated in observer performance testing. Data were analyzed in August 2018.

EXPOSURES: Deep learning–based algorithm.

MAIN OUTCOMES AND MEASURES: Image-wise classification performances measured by area under the receiver operating characteristic curve; lesion-wise localization performances measured by area under the alternative free-response receiver operating characteristic curve.

RESULTS: The algorithm demonstrated a median (range) area under the curve of 0.979 (0.973–1.000) for image-wise classification and 0.972 (0.923–0.985) for lesion-wise localization; the algorithm demonstrated significantly higher performance than all 3 physician groups in both image-wise classification (0.983 vs 0.814–0.932; all $P < .005$) and lesion-wise localization (0.985 vs 0.781–0.907; all $P < .001$). Significant improvements in both image-wise classification (0.814–0.932 to 0.904–0.958; all $P < .005$) and lesion-wise localization (0.781–0.907 to 0.873–0.938; all $P < .001$) were observed in all 3 physician groups with assistance of the algorithm.

CONCLUSIONS AND RELEVANCE: The algorithm consistently outperformed physicians, including thoracic radiologists, in the discrimination of chest radiographs with major thoracic diseases.

Key Points

Question: Can a deep learning–based algorithm accurately discriminate abnormal chest radiograph results showing major thoracic diseases from normal chest radiograph results?

Findings: In this diagnostic study of 54 221 chest radiographs with normal findings and 35 613 with abnormal findings, the deep learning–based algorithm for discrimination of chest radiographs with pulmonary malignant neoplasms, active tuberculosis, pneumonia, or pneumothorax demonstrated excellent and consistent performance throughout 5 independent data sets. The algorithm outperformed physicians, including radiologists, and enhanced physician performance when used as a second reader.

Meaning: A deep learning–based algorithm may help improve diagnostic accuracy in reading chest radiographs and assist in prioritizing chest radiographs, thereby increasing workflow efficiency.

Supplemental content

Author affiliations and article information are listed at the end of this article.



ORIGINAL ARTICLE
LUNG IMAGING

Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs

Ju Gang Nam^{1,2}, Minchul Kim³, Jongchan Park³, Eui Jin Hwang^{1,2}, Jong Hyuk Lee^{1,2}, Jung Hee Hong^{1,2}, Jin Mo Goo^{1,2,4} and Chang Min Park^{1,2,4}

Affiliations: ¹Dept of Radiology, Seoul National University Hospital, Seoul, Republic of Korea; ²College of Medicine, Seoul National University, Seoul, Republic of Korea; ³Lunit Inc., Seoul, Republic of Korea; ⁴Institute of Radiation Medicine, Seoul National University Medical Research Center, Seoul, Republic of Korea.

Correspondence: Chang Min Park, Dept of Radiology and Institute of Radiation Medicine, Seoul National University College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea. E-mail: cnpark.morphus@gmail.com

@ERSpublications

A deep learning algorithm detecting 10 common abnormalities was trained with 146 717 images and showed excellent performance on chest radiographs, helping radiologists improve their performance and advance the reporting time for critical or urgent cases <https://bit.ly/3k8Z3P>

Cite this article as: Nam JG, Kim M, Park J, et al. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. *Eur Respir J* 2021; 57: 2003061 [<https://doi.org/10.1183/13993003.03061-2020>].

iRAIL 소개

기술이전, 제품 상용화



Lunit INSIGHT CXR
종류 X-ray 영상 분석 솔루션

AI로 보다 효율적인 판독을 경험하세요

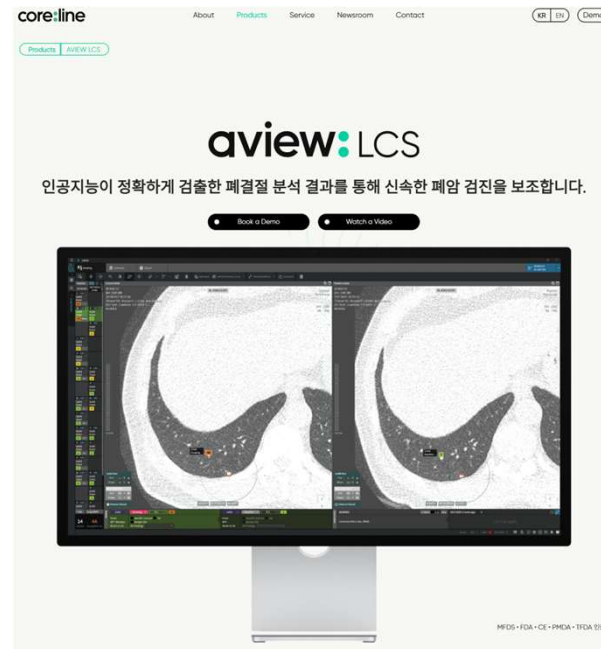
DICOM 파일을 업로드해 AI 분석 결과를 직접 확인해 보세요.

[온라인 데모 체험하기](#)

97-99% AUC
무엇보다도, CXR은 흉부 X-ray 영상에서 가장 흔하게 발견되는 90가지 이상 질환 소견을 97-99% 정확도로 검출합니다. (검체 선별 검사에도 사용할 수 있습니다.)

결핵	✓	결핵
기흉		흉수
무기폐		기흉증
상장폐대		중격동 확장
식회화		폐성유화

폐결핵 진단보조



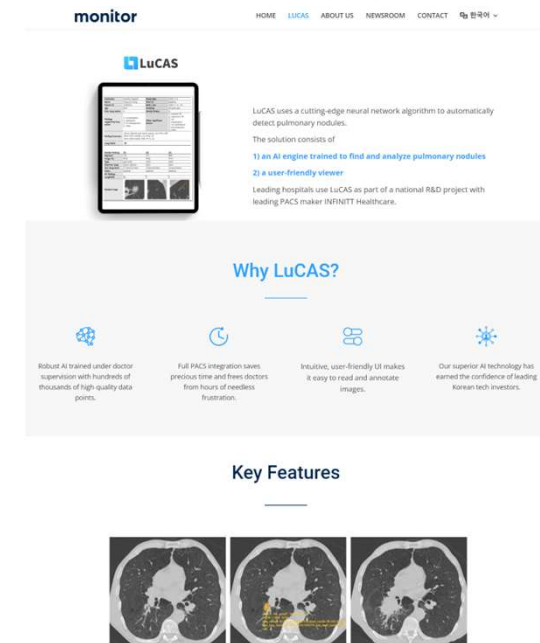
core:line

Products **aview:LCS**

인공지능이 정확하게 검출한 폐결절 분석 결과를 통해 신속한 폐암 검진을 보조합니다.

[Back a Demo](#) [Watch a Video](#)

MDTS - FDA - CE - PMDA - TFDA 인증



monitor

LuCAS

LuCAS uses a cutting-edge neural network algorithm to automatically detect pulmonary nodules. The solution consists of:

- 1) an AI engine trained to find and analyze pulmonary nodules
- 2) a user-friendly viewer

Leading hospitals use LuCAS as part of a national R&D project with leading PACS maker INFNITT Healthcare.

Why LuCAS?

- Robust AI trained under doctor supervision with hundreds of thousands of high-quality data points.
- Full PACS integration saves precious time and frees doctors from hours of needless frustration.
- Intuitive, user-friendly UI makes it easy to read and annotate images.
- Our superior AI technology has earned the confidence of leading Korean tech investors.

Key Features

 **Lunit**  **DRTECH**  **CORELINE**  **monitor**

임상 현장에서의 인공지능 교육에 대한 니즈 증가

2025 헬스케어 AI

연구인력 양성과정

2025.2.19

부터

2025.6.4

까지

교육기간
· 2025.2.19(수) ~ 2025.6.4(수)
매주 수요일 17시~19시

모집인원
· 50명

모집기간
· 2025.1.23(목) ~ 2025.1.30(목)

교육과정
· 서울대학교병원 김홍기원, 융합대학기술팀(서울 종로구)
· 서울 AI 허브(서울 서초구)

QR코드 접속 후 신청을 작성

지원자격

· 서울대학교병원, 분당서울대학교병원, 보라매병원, 연구중심병원 소속 임상사서 및 연구원

(연구중심병원 : 기초연구과, 노년과, 혈액내과, 신장내과, 피부과, 호흡기과, 구강악안면외과, 이비인후과, 소아청소년과, 응급의학과, 방사선과, 재활의학과, 감염내과, 소화기내과, 외과, 내과, 산부인과, 소아과, 진단의학과, 핵의학과(중성자)

· 서울대학교 대학원생 및 연구원

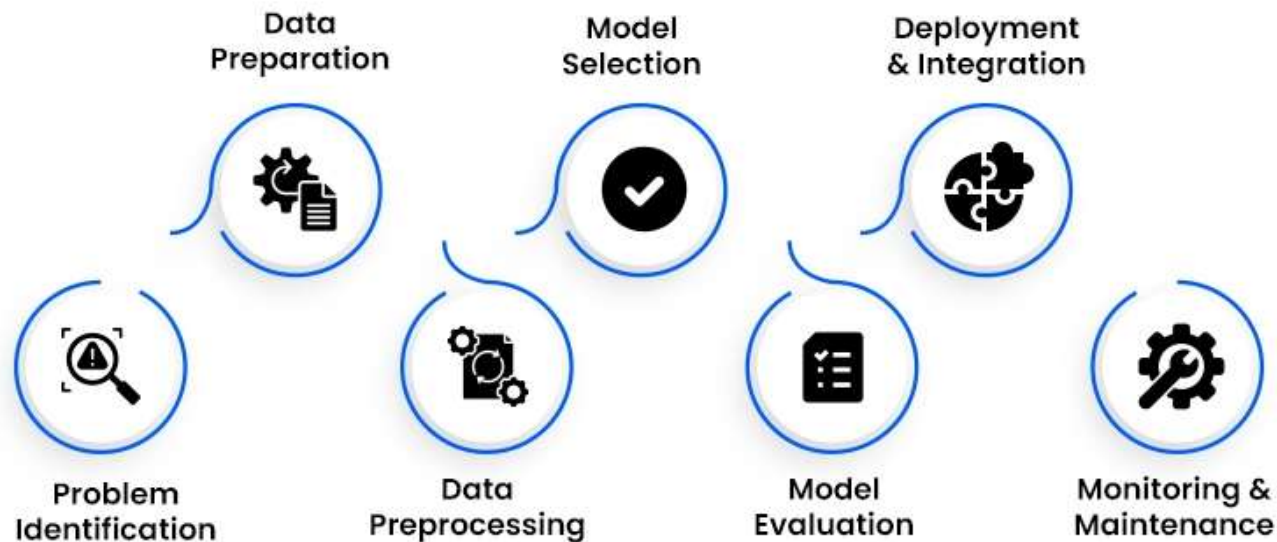
순위	날짜	수업 형태	내용
1순위	2월 19일	강의 / 실습	과정 소개 및 생체 인식 인공지능
2순위	2월 26일		유전체 분석을 위한 인공지능
3순위	3월 5일		의료 영상 분석의 인공지능
4순위	3월 12일		의료 분야 인공지능 이해
5순위	3월 19일		생체감각 인공지능 원리 이해
6순위	3월 26일		인공지능 분야 최신 이해
7순위	4월 2일		연구 학습실 원리 이해
8순위	4월 9일		의료 인공지능 연구 사례 소개
9순위	4월 16일	프로젝트	의료 인공지능 연구 사례 소개
10순위	4월 23일		의료 인공지능 연구 사례 소개
11순위	4월 30일		팀 프로젝트 및 주제 선정
12순위	5월 7일		팀 프로젝트 수행
13순위	5월 14일		중간 발표
14순위	5월 21일		팀 프로젝트 수행
15순위	5월 28일		팀 프로젝트 수행
16순위	6월 4일		최종 발표

[illegible]

의료 인공지능 연구의 bottle neck

Clinician의 아이디어가 실현 가능한지 파악하는 데에만 너무 많은 비용이 필요

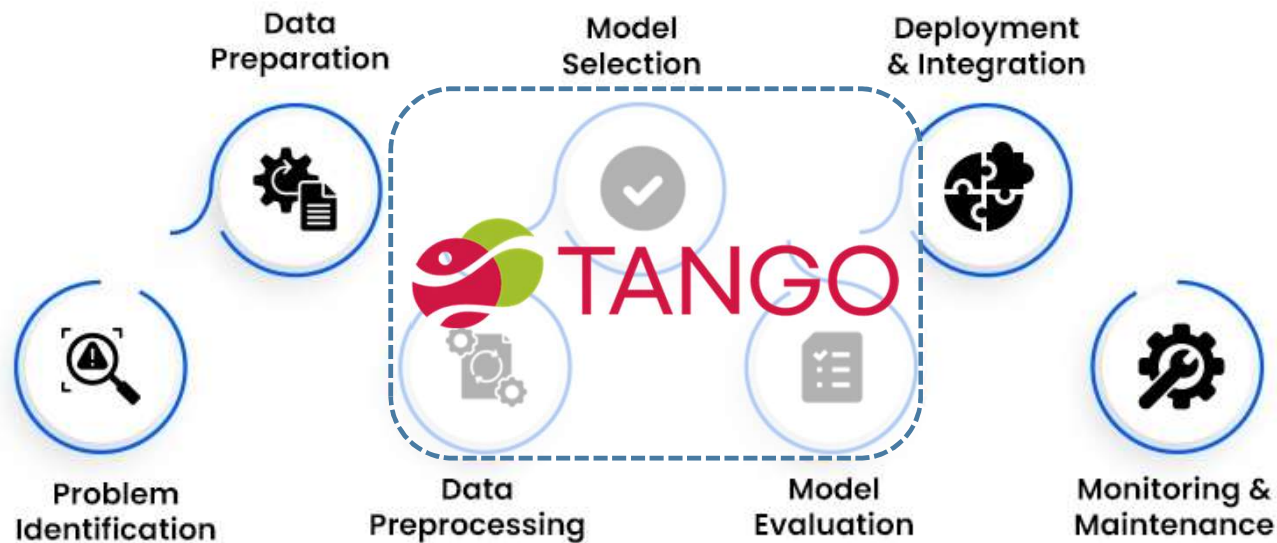
Key Stages of the AI Development Process



TANGO를 활용한 Proof-of-Concept

아이디어의 실현 가능성을 TANGO를 활용해 빠르게 검증, Proof-of-Concept(PoC)에 활용

Key Stages of the AI Development Process



활용 예시 1. 지역사회획득폐렴 환자의 흉부 X선 기반 딥러닝 모델 개발

지역사회획득폐렴

- 폐렴은 2021년 전세계 사망 240만 명(전체 3.6%)을 차지하는 주요 사망 원인 질병
- 고위험 환자 조기 입원 및 집중치료와 저위험 환자의 불필요한 입원과 침습적 치료 감소가 환자의 예후 개선 및 의료 비용 절감의 핵심
- 이와 관련하여 임상 지표를 기반으로 위험을 평가하는 지표가 연구되어져 왔음 (CURB-65, PSI)
- 흉부 X-ray는 폐렴 진단에 필수적이며 대부분의 환자에서 촬영되지만, 위험 예측에 영상 소견을 활용하는 것은 한계가 있음

→ 인공지능 기술을 활용해 지역사회획득폐렴 환자의 흉부 X-ray에서 30일 사망 위험을 예측하는 스 이으까?

Table 1. CURB-65 Scoring⁴⁶

Symptom	Points
Confusion	1
Urea: BUN >19 mg/dL (>7 mmol/L)	1
Respiratory rate ≥30 breaths/min	1
Systolic BP <90 mm Hg or diastolic BP ≤60 mm Hg	1
Age ≥65 years	1
Total	

Score	Risk	Disposition
0 or 1	1.5% mortality	Outpatient care
2	9.2% mortality	Inpatient versus observation admission
≥3	22% mortality	Inpatient admission; consider ICU admission with score of 4 or 5

Abbreviations: BP, blood pressure; BUN, blood urea nitrogen; ICU, intensive care unit.

PSI PNEUMONIA SEVERITY INDEX
FINE MJ ET AL: A PREDICTION RULE TO IDENTIFY LOW-RISK PATIENTS WITH
COMMUNITY ACQUIRED PNEUMONIA.NEJM 1997;336:243

Demographics	Co-morbidities	Physical exam / vital signs	Laboratory / imaging
<ul style="list-style-type: none"> • Age (1 point per year) • Male Yr +1 • Female Yr -10 • Nursing home residency +10 	<ul style="list-style-type: none"> • Neoplasia +30 • Liver disease +20 • CHF +10 • Cerebrovascular disease +10 • Renal disease +10 	<ul style="list-style-type: none"> • Mental confusion +20 • Respiratory rate +20 • SBP +20 • Temperature +15 • Tachycardia +15 	<ul style="list-style-type: none"> • Arterial pH +30 • BUN +20 • Sodium +20 • Glucose +10 • Hematocrit +10 • Pleural effusion +10 • Oxygenation +10

Risk class (Points)	Mortality (%)	Recommended site of care
I (<50)	0.1	Outpatient
II (51-70)	0.6	Outpatient
III (71-90)	2.8	Outpatient or brief inpatient
IV (91-130)	8.2	Inpatient
V (>130)	29.2	Inpatient

as a site-of-care tool. BUN, blood urea nitrogen; CHF, chronic heart failure; SBP, systolic blood pressure.

활용 예시 1. 지역사회획득폐렴 환자의 흉부 X선 기반 딥러닝 모델 개발

9

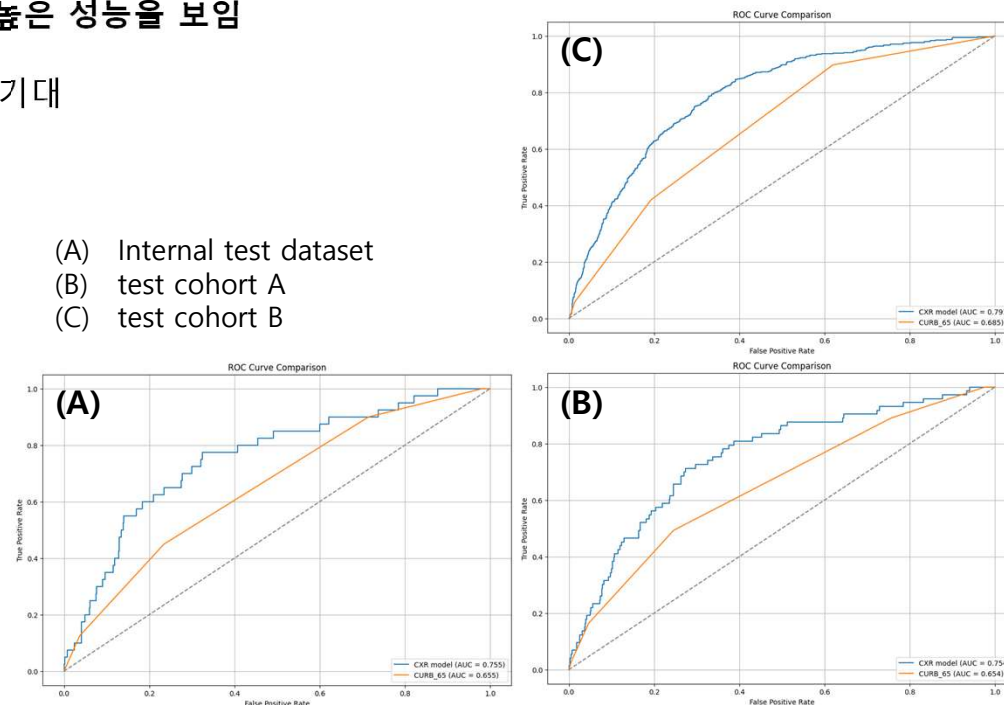
흉부 X선 기반 30일 사망률 예측 모델 개발

- 서울대병원 응급실에 내원한 폐렴 환자의 흉부 X선 데이터를 활용해 30일 사망 위험을 예측하는 딥러닝 모델 개발
- 기존의 위험 예측 지표인 CURB-65와 비교했을 때 유의하게 더 높은 성능을 보임

→ 폐렴 환자의 사망 위험 예측을 통한 치료 방침 결정 도구로 활용 기대

	Internal test dataset (n=696)	Test cohort A (n=996)	Test cohort B (n=4,703)
Area under receiver operating characteristic curve			
CXR model	0.76 (0.68–0.83)	0.75 (0.70–0.81)	0.79 (0.77–0.81)
CURB-65	0.65 (0.57–0.73)	0.66 (0.59–0.72)	0.69 (0.66–0.71)
Harrell's C-index			
CXR model	0.69 (0.65–0.73)	0.66 (0.63–0.69)	0.74 (0.72–0.75)
CURB-65	0.61 (0.57–0.65)	0.60 (0.56–0.63)	0.66 (0.64–0.67)

- (A) Internal test dataset
(B) test cohort A
(C) test cohort B



활용 예시 2. 건강검진 수검자의 6년 이내 폐암 발생 위험 예측을 위한 흉부 X선 기반 딥러닝 모델 개발

폐암

- 폐암은 여러 암종 중 사망률 1위 질환으로 **전체 암사망 20% 이상**을 차지하고 있음
- 2019년부터 국가암검진사업에 **폐암 고위험군 대상 저선량 흉부 CT 검사**가 추가됨
- 고위험군 현행 기준(연령, 흡연력)으로는 **고위험 대상 누락과 불필요한 검진 발생**

암종	검진대상	검진주기	검진방법
 위암	40세 이상 남녀	2년	위내시경검사
 간암	40세 이상 남녀 간암발생고위험군 (간경변증이나 B형 간염 바이러스 항원 또는 C형 간염바이러스 항체 양성으로 확인된 자)	6개월	간초음파검사 + 혈청알파태단백검사
 대장암	50세 이상 남녀	1년	분변잠혈검사 : 이상소견시 대장내시경검사
 유방암	40세 이상 여성	2년	유방촬영술
 자궁경부암	20세 이상 여성	2년	자궁경부세포검사
 폐암	54세 이상 74세 이하의 남·여 중 폐암 발생 고위험군	2년	저선량흉부 CT

폐암검진

54세~74세 남녀 중 고위험군
2년마다

저선량흉부CT 검사



폐암검진 대상 고위험군 기준

30갑년 [하루 평균 담배소비량(갑)x흡연기간(년)] 이상의 흡연력을 가진 흡연자

- 해당연도 전 2년내 일반건강검진(생애전환기 건강진단 포함)의 문진표로 흡연력과 현재 흡연 여부가 확인되는 자

- 해당연도 전 2년내 건강보험 금연치료 참여자 중 사업참여를 위해 작성하는 문진표로 흡연력이 확인되는 자

폐암 발생 고위험군으로 확인되어 국가폐암검진을 받았던 자는 검진 후 금연을 하더라도 금연 15년 이내, 74세까지 폐암 검진 대상자에 포함

활용 예시 2. 건강검진 수검자의 6년 이내 폐암 발생 위험 예측을 위한 흉부 X선 기반 딥러닝 모델 개발

폐암 위험도 예측 모델 사례 및 흉부 X선 영상의 활용 가능성

- 폐암 고위험군 대상자 선별의 중요도가 높아 연령과 흡연력 이외 다양한 폐암의 위험인자를 반영한 위험 예측 모델 연구가 꾸준히 이루어짐
- PLCOall2014**: 임상 정보를 바탕으로 6년 내 폐암 발생 위험을 예측하는 다변량 모델 (흡연자, 비흡연자 모두 적용)
- 흉부 X선 검사는 폐암 예측에 필요한 정보를 내포하고 있을 것으로 기대되며, 일반적으로 가장 자주 시행되는 영상 검사임을 고려할 때 예측 도구로서 활용도 높을 것으로 예상됨

→ 인공지능 기술을 활용해 건강검진 수검자의 흉부 X-ray에서 6년 이내 폐암 발생 위험을 예측할 수 있을까?

Variables	Model PLCO participants sampled	PLCO _{all2012} † Ever-smokers only	PLCO _{all2014} § Never- & ever-smokers
		Odds Ratio (95% CI, p-value)	Odds Ratio (95% CI, p-value)
Age † (per one year)		1.081 (1.057 to 1.105; p < 0.001)	1.083 (1.061 to 1.105; p < 0.001)
Race/ethnicity (self-reported)			
White		Referent group	Referent group
Black		1.484 (1.083 to 2.033; p = 0.014)	1.379 (1.009 to 1.885; p = 0.044)
Hispanic		0.475 (0.195 to 1.160; p = 0.102)	0.440 (0.181 to 1.073; p = 0.071)
Asian		0.627 (0.332 to 1.185; p = 0.150)	0.592 (0.323 to 1.085; p = 0.090)
Native Hawaiian or Pacific Islander		1	0.256 (0.036 to 1.838; p = 0.175)
American Indian or Alaskan Native		2.793 (0.992 to 7.862; p = 0.052)	2.591 (0.926 to 7.248; p = 0.070)
Education † (per 6 levels)		0.922 (0.874 to 0.972; p = 0.003)	0.916 (0.870 to 0.964; p = 0.001)
Body mass index † (weight in kilograms/height in meters squared)		0.973 (0.955 to 0.991; p = 0.003)	0.971 (0.954 to 0.988; p = 0.001)
Chronic obstructive pulmonary disease (yes vs. no)		1.427 (1.162 to 1.754; p = 0.001)	1.413 (1.153 to 1.730; p = 0.001)
Personal history of cancer (yes vs. no)		1.582 (1.172 to 2.128; p = 0.003)	1.623 (1.219 to 2.162; p = 0.001)
Family history of lung cancer (yes vs. no)		1.799 (1.471 to 2.200; p < 0.001)	1.796 (1.479 to 2.182; p < 0.001)
Smoking status (former=0, current=1)		1.297 (1.047 to 1.605; p = 0.017)	Referent group
Former smoker (yes vs. no)		Not applicable	10.082 (7.060 to 14.397; p < 0.001)
Current smoker (yes vs. no)		Not applicable	17.995 (11.276 to 28.710; p < 0.001)
Smoking intensity (average cigarettes/day)		p < 0.001 ‡	p < 0.001 §
Smoking duration (per year)		1.032 (1.014 to 1.051; p = 0.001)	1.031 (1.013 to 1.049; p = 0.001)
Smoking quit-time		0.970 (0.950 to 0.990; p = 0.003)	0.968 (0.949 to 0.988; p = 0.002)
Model constant		-4.532506 (unexponentiated)	-6.84088 (unexponentiated)
ROC AUC (95% CI)			
AUC in PLCO control (development) arm (ever- & never-smokers)		0.803 (0.782 to 0.813) (smokers only, n=36286)	0.859 (0.845 to 0.872) (all, n=68706)
AUC in PLCO intervention (validation) arm (ever- & never-smokers)		0.797 (0.782 to 0.813) (smokers only, n=37,332)	0.848 (0.833 to 0.861) (all, n=70540)
AUC in PLCO never-smokers in control & intervention arms		Not applicable	0.662 (0.607 to 0.709) (n=65711)
Predictive Performance – Calibration – Cox Recalibration in PLCO Intervention Arm ¶			
Recalibrated factor added to adjust the original model intercept		-0.131 (-0.394 to 0.132; p = 0.14)	-0.138 (-0.361 to 0.084; p = 0.223)
Recalibration factor to multiply the original model logit		0.946 (0.874 to 1.018; p = 0.33)	0.944 (0.886 to 1.000; p = 0.0502)
Predictive Performance – Calibration – Median & 90 th percentile absolute error			
PLCO control (development) arm (ever- & never-smokers)		0.0087, 0.04	0.0002, 0.0005
PLCO intervention (validation) arm (ever- & never-smokers)		0.0092, 0.04	0.0003, 0.0014
PLCO never-smokers in control & intervention arms		Not applicable	0.0002, 0.0003

Clinical variables: age, race, education, BMI, COPD, personal history of cancer, smoking status, smoking intensity, smoking duration, smoking quit-time,

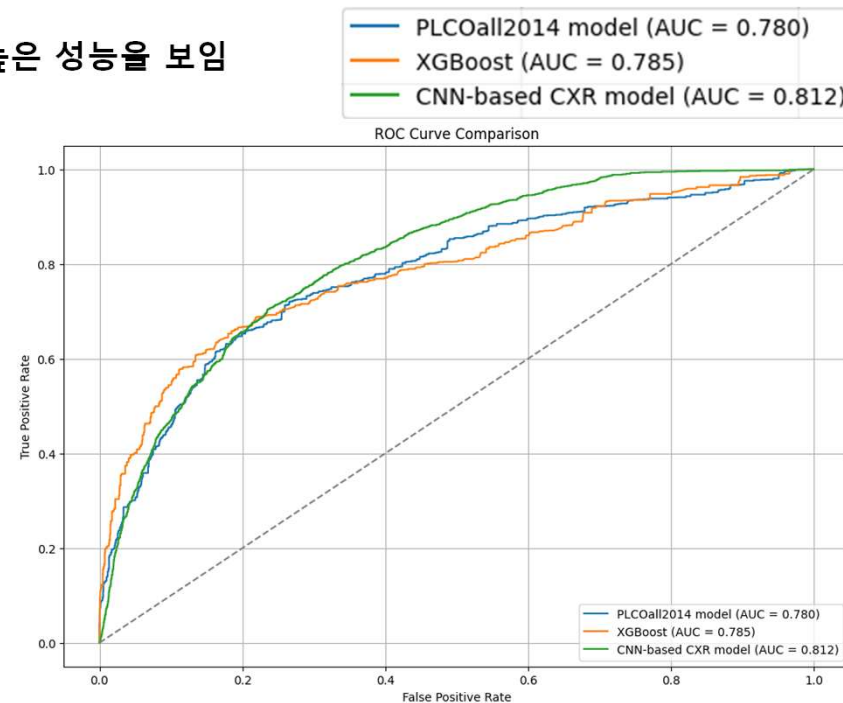
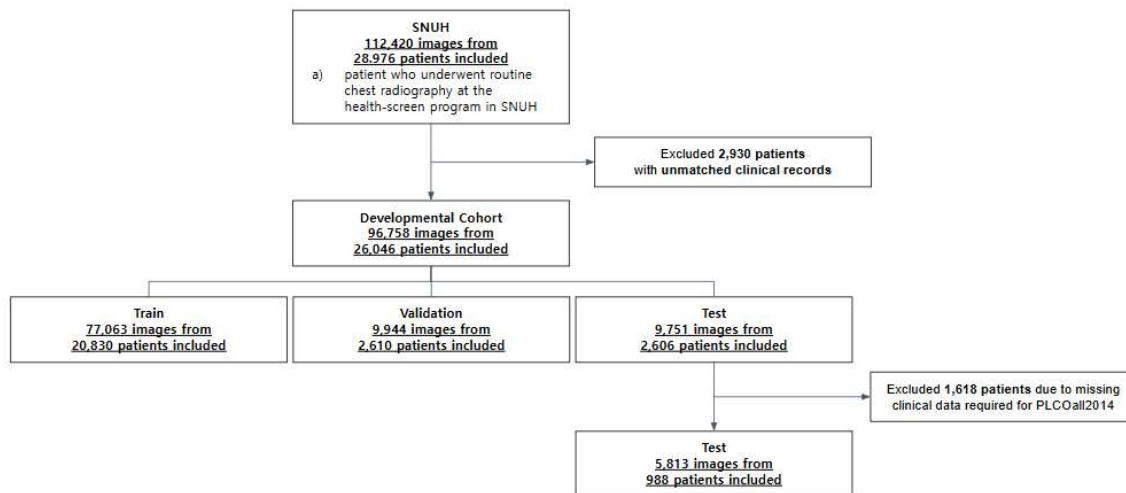
활용 예시 2. 건강검진 수검자의 6년 이내 폐암 발생 위험 예측을 위한 흉부 X선 기반 딥러닝 모델 개발

12

흉부 X선 기반 폐암 발생 위험 예측 모델 개발

- 서울대병원 건강증진센터 건강검진 수검자의 흉부 X선 데이터를 활용해 6년 이내 폐암 발생 위험을 예측하는 딥러닝 모델 개발
- 기존의 위험 예측 지표인 PLCOall2014 모델과 비교했을 때 유의하게 더 높은 성능을 보임

→ 폐암 발생 위험 예측을 통한 폐암검진 대상자 선정의 최적화 기대

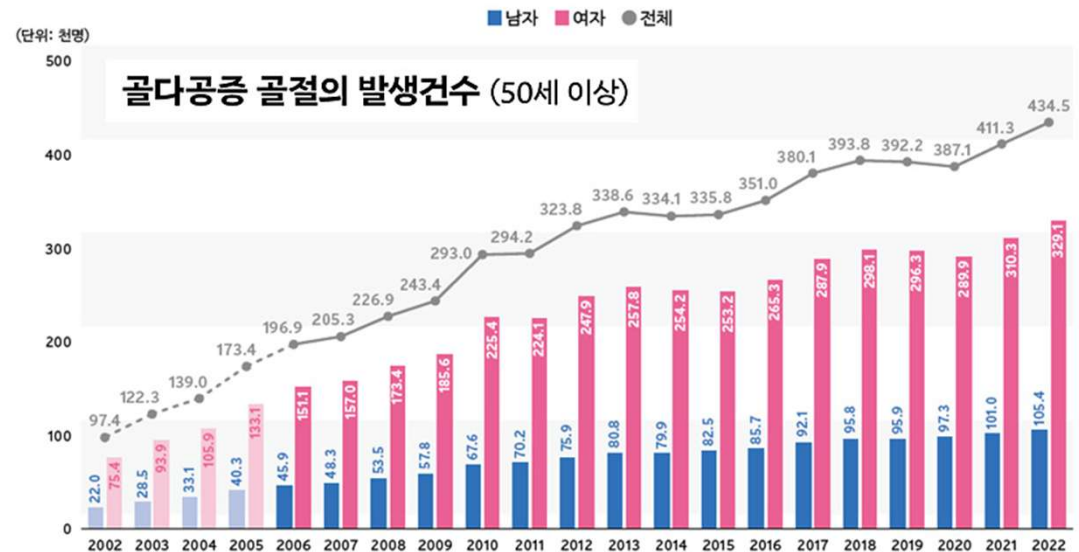
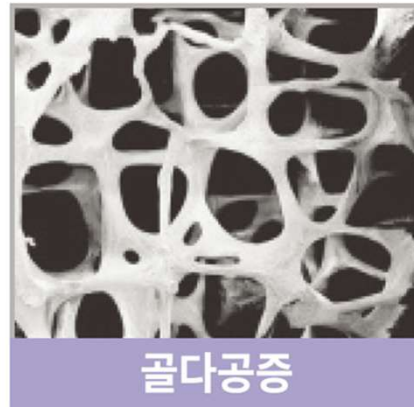


활용 예시 3. 골밀도 검사자의 골다공증성 골절 발생 위험 예측을 위한

13

골다공증성 골절

- 골다공증성 골절은 넘어지는 정도의 약한 충격으로도 발생하는 골절
- 고령화 사회로 진행하면서 환자 증가, **연간 1조 5천억원 규모의 사회적 비용 소요**
- **50세 이상 여성 10명 중 3~4명 골절 위험군이지만 선별에 특수검사 필요해 진단 및 치료 시작 어려움**



활용 예시 3. 골밀도 검사자의 골다공증성 골절 발생 위험 예측을 위한

14

기존 골절 예측 도구: 골밀도 검사와 FRAX

- 예방을 위해 위험군 선별하여 치료를 수행하는 것이 중요
 - 골밀도 검사 촬영으로 골밀도를 파악해 FRAX로 위험도 예측
 - 골밀도 검사 촬영은 건강 검진 목적으로 시행 받는 환자군 제한적, 기회검진 가능성 낮음
- 흔히 시행되는 흉부 X-ray를 인공지능으로 분석해 골다공증성 골절 발생 위험을 예측할 수 있지 않을까?

골절 위험도 계산

골밀도를 알고 있는 경우 10년 내 골절 위험도 계산을 위하여 아래의 질문에 대하여 주십시오.

국가: 영국 이름 / 아이디: 위험 요인에 대하여

질문 사항:

1. 출생의 시대 (40-90년 사이) 또는 날짜
나이: 생년월일: 년: 월: 주:

2. 성별 ☐ 남성 ☐ 여성

3. 체중 (kg)

4. 신장 (cm)

5. 이전의 골절 병력 ☒ 아니오 ☐ 예

6. 부모님의 고관절 골절 ☒ 아니오 ☐ 예

7. 현재 흡연 ☒ 아니오 ☐ 예

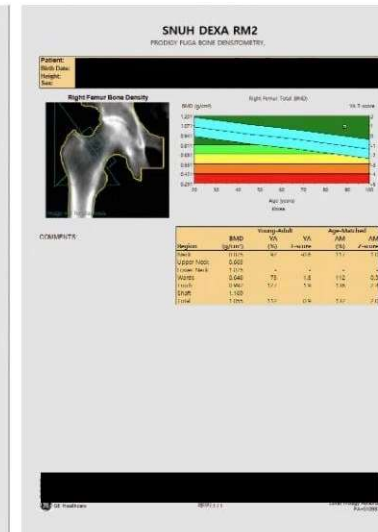
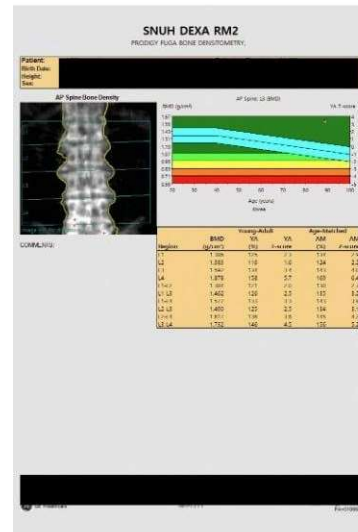
8. 스테로이드 제제 ☒ 아니오 ☐ 예

9. 류마티스 관절염 ☒ 아니오 ☐ 예

10. 이차성 골다공증 ☒ 아니오 ☐ 예

11. 하루 3단위 이상의 술 ☒ 아니오 ☐ 예

12. 대퇴골 경부 (BMD) (g/cm2)
선택 BMD



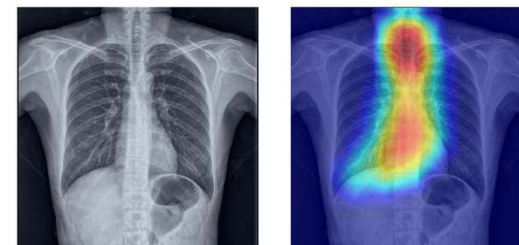
활용 예시 3. 골밀도 검사자의 골다공증성 골절 발생 위험 예측을 위한

15

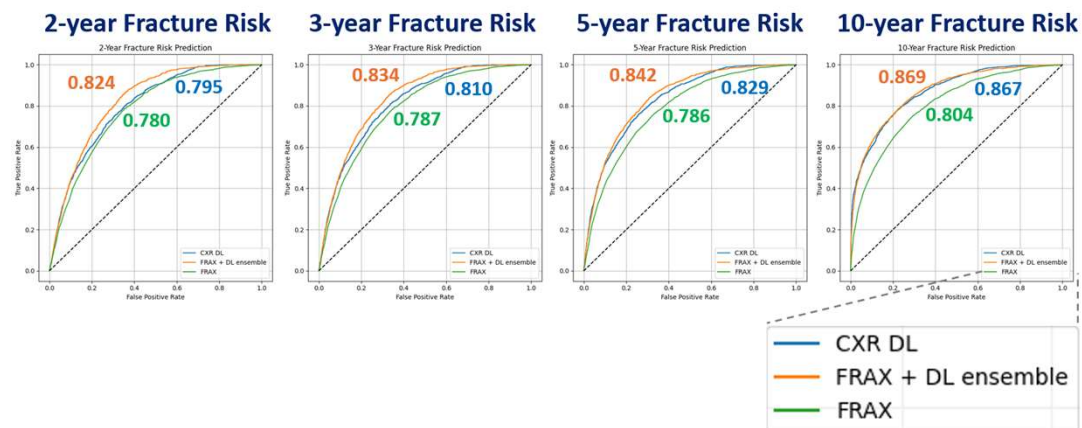
흉부 X선 기반 딥러닝 모델 개발

- 서울대병원 흉부 X선 데이터를 활용해 골다공증성 골절 발생 위험을 예측하는 딥러닝 모델 개발
- 기존에 임상에서 활용되던 **FRAX**에 비해 **유의하게 더 높은 성능** 보이는 것을 확인

→ 건강검진 등에서 촬영한 흉부 X선 활용해 골절 위험군 미리 선별 및 치료 가능할 것으로 기대



	C-index	2y-AUROC	3y-AUROC	5y-AUROC	10y-AUROC
Internal test set (n=5,000)					
CXR DL	0.867±0.012*	0.878±0.016*	0.887±0.014*	0.886±0.014*	0.899±0.014*
FRAX + DL ensemble	0.859±0.014*	0.869±0.016*	0.873±0.014*	0.874±0.012*	0.892±0.012*
FRAX	0.800±0.016	0.805±0.024	0.804±0.022	0.805±0.020	0.832±0.020
External test sets (n=37,865)					
CXR DL	0.788±0.006*	0.795±0.006*	0.810±0.006*	0.829±0.007*	0.867±0.007*
FRAX + DL ensemble	0.811±0.006*	0.824±0.006*	0.834±0.006*	0.842±0.007*	0.869±0.007*
FRAX	0.772±0.006	0.780±0.007	0.787±0.007	0.786±0.007	0.804±0.006



감사합니다.



주 관 ETRI (TANGO)
주 최 과학기술정보통신부 IITP 정보통신기획평가원
문 의 parkjb@etri.re.kr / 042-860-5565

후 원 LG labup w e o a tesla system (사)한국인공지능협회 SNUH 서울과학기술대 고려대학교 KOREA UNIVERSITY 영인대학교 YOUNG IN UNIVERSITY 중앙대학교 CHUNG ANG UNIVERSITY RTst Reliable & Trustworthy

KEITI 한국전자기술연구원 AIVN 한국인공지능협회 SUREDATA ACRYL h 하일소프트 한국정보통신기술협회