ISSN 2644-3260 (online)

Artificial Intelligence in *Medical Imaging*

Artif Intell Med Imaging 2022 April 28; 3(2): 21-54

Published by Baishideng Publishing Group Inc

AIMI

Artificial Intelligence in Medical Imaging

Contents

Bimonthly Volume 3 Number 2 April 28, 2022

MINIREVIEWS

SYSTEMATIC REVIEWS

Artificial Intelligence in Medical Imaging

Bimonthly Volume 3 Number 2 April 28, 2022

Contents

ABOUT COVER

Editorial Board Member of *Artificial Intelligence in Medical Imaging*, Drozdstoy Stoyanov, MD, PhD, DSc, IDFAPA, Professor, Head Department of Psychiatry and Medical Psychology, Scientific Secretary of the Research Institute at Medical University Plovdiv; Vice President of the Bulgarian Academy of Sciences and Arts; Vassil Aprilov 15 a, Plovdiv 4002, Bulgaria. drozdstoy.stoyanov@mu-plovdiv.bg

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Medical Imaging* (*AIMI*, *Artif Intell Med Imaging*) is to provide scholars and readers from various fields of artificial intelligence in medical imaging with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIMI mainly publishes articles reporting research results obtained in the field of artificial intelligence in medical imaging and covering a wide range of topics, including artificial intelligence in radiology, pathology image analysis, endoscopy, molecular imaging, and ultrasonography.

INDEXING/ABSTRACTING

The *AIMI* is now abstracted and indexed in Reference Citation Analysis, China Science and Technology Journal Database.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Ying-Yi Yuan*; **Production Department Director:** *Xu Guo*; **Editorial Office Director:** *Yun-Xiaojiao Wu*.

NAME OF JOURNAL
Artificial Intelligence in Medical Imaging

ISSN
ISSN 2644-3260 (online)

LAUNCH DATE
June 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Xue-Li Chen, Caroline Chung, Jun Shen

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2644-3260/editorialboard.htm

PUBLICATION DATE
April 28, 2022

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

Artificial Intelligence in Medical Imaging

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Med Imaging 2022 April 28; 3(2): 21-32

DOI: 10.35711/aimi.v3.i2.21

ISSN 2644-3260 (online)

MINIREVIEWS

Role of artificial intelligence in early detection and screening for pancreatic adenocarcinoma

Kenneth Weicong Lin, Tiing Leong Ang, James Weiquan Li

Specialty type: Gastroenterology and hepatology

Provenance and peer review: Invited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

P-Reviewer: Nazari N, Iran; Zhang JX, China

Received: December 16, 2021
Peer-review started: December 16, 2021
First decision: January 26, 2022
Revised: February 12, 2022
Accepted: March 16, 2022
Article in press: March 16, 2022
Published online: April 28, 2022

Kenneth Weicong Lin, Tiing Leong Ang, James Weiquan Li, Department of Gastroenterology and Hepatology, Changi General Hospital, Singapore 529889, Singapore

Corresponding author: James Weiquan Li, FRCPE, MBBS, MMed, Assistant Professor, Doctor, Department of Gastroenterology and Hepatology, Changi General Hospital, 2 Simei Street 3, Singapore 529889, Singapore. james.li.w.q@singhealth.com.sg

Abstract

Pancreatic adenocarcinoma remains to be one of the deadliest malignancies in the world despite treatment advancement over the past few decades. Its low survival rates and poor prognosis can be attributed to ambiguity in recommendations for screening and late symptom onset, contributing to its late presentation. In the recent years, artificial intelligence (AI) as emerged as a field to aid in the process of clinical decision making. Considerable efforts have been made in the realm of AI to screen for and predict future development of pancreatic ductal adenocarcinoma. This review discusses the use of AI in early detection and screening for pancreatic adenocarcinoma, and factors which may limit its use in a clinical setting.

Key Words: Artificial intelligence; Pancreatic cancer; Pancreatic adenocarcinoma; screening; Early detection

Core Tip: Pancreatic adenocarcinoma has poor survival rate and high morbidity. Artificial intelligence is a potential tool to screen for high risk individuals and for early detection of pancreatic adenocarcinoma. Despite advances made in artificial intelligence research in pancreatic adenocarcinoma, it faces a number of challenges before it can be generalised and applied in a clinical setting.

Citation: Lin KW, Ang TL, Li JW. Role of artificial intelligence in early detection and screening for pancreatic adenocarcinoma. *Artif Intell Med Imaging* 2022; 3(2): 21-32
URL: https://www.wjgnet.com/2644-3260/full/v3/i2/21.htm
DOI: https://dx.doi.org/10.35711/aimi.v3.i2.21

INTRODUCTION

The global incidence of pancreatic cancer is increasing, and it remains as one of the leading causes of cancer-related death, with 495773 new cases of pancreatic cancer diagnosed and accounting for 466003 deaths in 2020[1]. Although the 5-year survival rates for pancreatic ductal adenocarcinoma (PDAC) have improved, it remains low at approximately 9%[2,3], and the overall prognosis of PDAC is poor. This is partly due to the late stage of presentation of PDAC, which is largely dependent on patient symptoms for suspicion of the disease[4,5]. Early cases are asymptomatic and there is a lack of a simple screening tool for clinical use unlike the case of colorectal cancer screening where screening can be performed in the primary care setting with the use of fecal immunohistochemical test. In the case of PDAC, cross-sectional imaging tests such as computed tomography (CT) or magnetic resonance imaging (MRI) are needed for detection, making widespread population screening unfeasible. Germline mutations and a family history of PDAC have been identified as the strongest risk factors for the disease [6,7]. As such, efforts in screening programmes have focused their attention on this group of patients[8]. Pancreatic cysts, increased age, and smoking are also known risk factors for PDAC[5,9,10], although it may not be practical to conduct routine surveillance for patients with these risk factors. There is an interest in selecting higher risk patients for screening, as the appropriate use biomarkers and imaging may result in detection of early-stage PDAC amenable to curative resection[2,3,11-15].

Artificial intelligence (AI) is a branch in computer science where computer systems are designed to perform tasks which would require human intelligence. It is recognised as a potential tool as part of the screening efforts and building predictive models[16]. Most progress for AI in endoscopy has been made in the field of colonoscopy, where polyp detection and characterisation has been studied[17]. Computer-aided diagnosis has also been extended to detection and screening of PDAC[18] in endoscopic ultrasound (EUS)[19,20], MRI[21] and cytology from fine needle sampling[22]. In recent years, various groups have harnessed the potential of AI in creating prediction models. These include The Felix Project [23], the Pancreatic-Cancer Collective[24], and the Early Detection Research Network[25] effort.

This mini-review aims to study the role of AI in the early detection and screening for pancreatic cancer, as well as factors which may limit its use.

METHODS

A comprehensive literature search was performed in the PubMed, MEDLINE and EMBASE electronic databases from the inception of the databases up to and including 30 November 2021. The key words used were "artificial intelligence", "pancreatic cancer", "pancreatic adenocarcinoma", "pancreatic ductal adenocarcinoma", "pancreatic carcinoma", "screening", and "early detection". These were supplemented with manual searches of references from retrieved articles. Publications in English were considered for this mini-review.

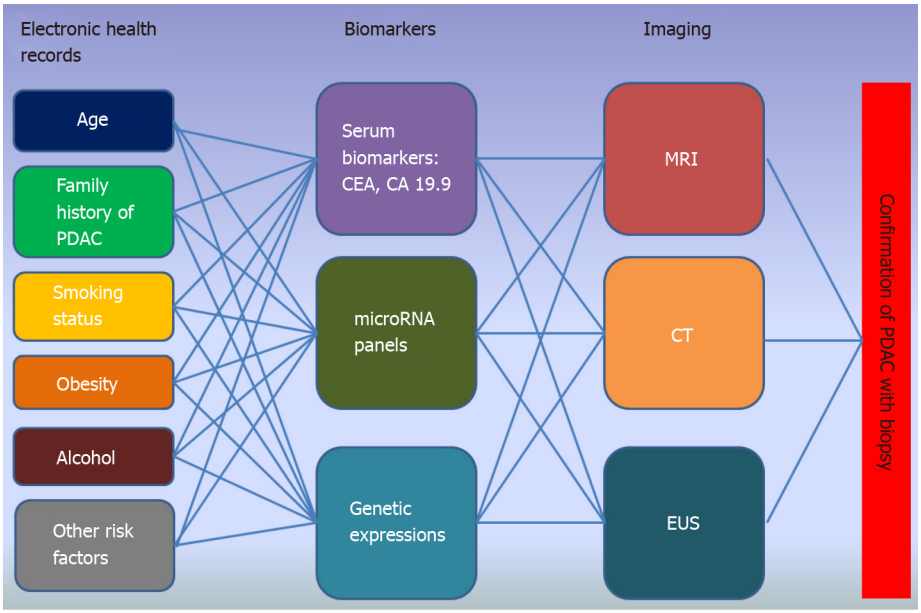
AI BASIC PRINCIPLES AND TERMINOLOGIES

AI is a term that refers to the ability of a computer programme to imitate the human mind to perform tasks such as problem solving and learning[26,27].

Machine learning (ML) is the commonest branch of AI used in medicine and refers to a mathematical model that aims to generate a prediction based on a set of data provided[28,29]. In supervised learning, the data points are labelled and the ML model "learns" from these labels and identifies new data points. In contrast, labels are not provided in unsupervised learning, and the model recognises the patterns of the data by learning its unknown properties and identifying crucial data checkpoints. This is especially important when the gold standard is not available[29].

Deep learning (DL) is subset of ML that employs the use of Artificial Neural Networks (ANN). Like the human brain, ANN consists of layers of artificial neurons that are interlinked. Each layer receives a weighted signal from the previous layer(s) and these signals will be propagated to the next layer when a specific threshold is exceeded[29]. In the setting of a pancreatic lesion or cancer, DL first identifies the basics of the lesion (*e.g.*, location) in its initial layers before moving on to next layer for further characterisation (*e.g.*, size, shape, colour). A final prediction of the pancreatic lesion is made after a systematic assessment *via* multiple layers of neural network[29].

ANNs are first trained using the training data set, where the model learns to identify specific patterns to obtain a relationship between the input and the output. Hyperparameters refer to all settings that are pre-determined by the investigator and are used to construct the model for optimal execution of a particular task or on a specific dataset. The validation data set involves a different data set that is used to fine-tune the hyperparameters of the model. Finally, the test data set refers to a data set whose purpose is to evaluate the performance of the model against unseen data and determine its generalizability[29]. This set needs to be unseen by the model during training and validation. However in certain studies, the test set is sometimes a subset of the training or validation data set, which many result in



DOI: 10.35711/aimi.v3.i2.21

Figure 1 Schematic diagram showing the workflow and neural network to be designed for an early detection protocol. CT: Computed tomography; CEA: Carcinoembryonic antigen; PDAC: Pancreatic ductal adenocarcinoma; MRI: Magnetic resonance imaging.

overfitting of the model. This may lead to a discrepancy in the performance of the model when tested in the same centre and a decline in performance when validated externally.

MODEL FOR SCREENING FOR AND EARLY IDENFICATION OF DEVELOPING PDAC

Early detection of pancreatic cancer requires a step wise approach in order to systematically screen for risk factors and identify high-risk groups. Figure 1 is a schematic diagram showing the workflow and neural network to be designed for an early detection protocol. It represents the complex interplay between each of the input(s) to be processed for the next neural layer(s) until a final output is obtained. We will be discussing the role of AI in early detection of pancreatic cancer based on this model.

AI IN CLINICAL DECISION MAKING USING HEALTH RECORDS

The identification of risk factors for pancreatic cancer is essential in identifying the specific population which would benefit from screening[18,30,31]. Factors such as diabetes, hemoglobin A1C (HbA1c) value, weight, body mass index (BMI), blood type, smoking status, alcohol use and family history of pancreatic cancer influence the age of onset of screening for an individual[13,32]. These factors are easily available in the primary care setting and could potentially predict the development of pancreatic cancer within 5 years, even before any changes to the pancreas can be detected on imaging[30]. However, most of the data is stored in health records, which are often proprietary or internet-separated to protect patient data. The retrieval and subsequent integration of data from different platforms remains a manual and laborious process for physicians[30]. Even after retrieval, there are no validated scoring systems to assess these risk factors and stratify patients. On the other hand, AI, with the aid of Natural Language Processing, can facilitate this process[33-38]. In a case-control study, Malhotra *et al*[33] created an algorithm based on electronic health records (EHR) obtained from primary care to identify 41.3% of patients (≤ 60 years old) who had significant risk of developing pancreatic cancer up to 20 mo prior to diagnosis with a sensitivity, specificity, area under the receiver operating characteristic (AUROC) curve of 72.5%, 59.0% and 0.66%, respectively. Similarly, Appelbaum *et al*[35] was able to train an ANN using 101381 EHRs to predict the development of PDAC one year before the diagnosis in a population of high-risk patients (AUROC 0.68, confidence interval (CI): 0.65-0.71).

Despite its potential benefits, research in AI for the above purpose is still preliminary as they are mostly based on retrospective data from single institutions or registries, and hence not ready for use in a wider clinical setting[33-38]. One of the major limitations would be the lack validation in the real-world setting or at least in populations derived from different centres to overcome the risk of bias and overfitting.

The use of AI in EHR faces other challenges. Various institutions' medical records are built on different healthcare systems and encoding systems, making the task of harmonising them difficult[30]. There is also a lack of standardised clinical research data collection models. To overcome this, efforts are made to build a model of processing and integrating data across institutions. The i2b2 was created to review medical records, retrieve specific data of interest and repurpose it for research[39]. The Observational Health Data Sciences and Informatics was developed from the Observational Medical Outcomes Partnership, an initiative that develops the Common Data Model aiming to gather information from different data sets or medical repositories and systemically analyse them in a common platform[40]. Similarly, the National Patient-centered Clinical research network is another example which was developed in United States to access millions of EHR and create a common data set for research purposes[41]. A common dataset with a standardised format for input of data relevant to PDAC would enable AI systems to leverage on big data to identify changing risk profiles in PDAC, enabling the clinician to channel resources for screening to the appropriate cohorts of patients depending on the population from which this data has been derived.

While these are upcoming and promising initiatives, concerns surrounding restrictions in data sharing, privacy issues, and maintenance costs could hinder data collection efforts[18]. EHRs are also stored in different languages in different regions of the world, making the integration of data difficult. Besides, once data sets are gathered, obtaining IRB approval from the various sites for research may be difficult.

AI AND THE USE OF NON-INVASIVE BIOMARKERS

Carbohydrate Antigen 19-9 (CA19-9) and carcinoembryonic antigen (CEA) are the most widely used markers for screening of PDAC, but have also been proven to lack the specificity when applied individually and without clinical context[42,43]. On the other hand, a combined measurement can potentially increase its sensitivity and specificity up to 1 year before the diagnosis of PDAC[44-46]. Capitalising on this concept, Yang *et al*[47], developed an algorithm (with 658 subjects in its training set) to diagnose pancreatic cancer by using ANN to combine CA19-9, CA125 and CEA values. This model was subsequently evaluated against the test set and was able to yield an AUROC of 0.905 (95%CI = 0.868-0.942) and a high diagnostic accuracy of 83.5% for pancreatic cancer.

New biomarkers for PDAC such as MicroRNAs and gene expressions have generated much interest in the recent years[45,48-52]. MicroRNAs are non-coding RNAs that are involved in the regulation of biological pathways, and when altered, could lead to the development of PDAC[53]. MicroRNAs can potentially predict future PDAC[54] or detect early stage pancreatic cancer. However, they have the same limitations in sensitivity and specificity when applied without clinical context and as independent test[55,56]. A combination of the commonly used biomarkers and newer biomarkers may address the problem of low sensitivity and specificity[56], and in particular can be combined with clinical and demographic information as described earlier to increase its usefulness.

While AI is able to make use of plasma microRNA panels and specific gene expressions to diagnose pancreatic cancer[57,58], studies on their use on predicting future pancreatic cancer are not available [55]. By integrating Particle Swarm Optimization, ANN and Neighborhood Component Analysis iterations on a list of microRNAs that are most commonly expressed by pancreatic cancer, Alizadeh *et al* [59] created a model consisting of 5 MicroRNAs (miR-663a, miR-1469, miR-92a-2-5p, miR-125b-1-3p and miR-532-5p) to diagnose pancreatic cancer (Accuracy: 0.93, Sensitivity: 93%, and Specificity: 92%). Similarly in a multicentre study by Cao *et al*[57], a machine learning approach was able to identify 2 panels of microRNAs to differentiate pancreatic cancer from chronic pancreatitis with an accuracy of above 80%.

Gene expressions have gained popularity in diagnosing pancreatic cancer[13,60]. Using a machine learning approach, Khatri *et al*[61] analysed the results from transcriptomics-based meta-analysis to create a nine-gene panel to diagnose pancreatic cancer. This panel was able to differentiate PDAC from chronic pancreatitis with a specificity of 89%, sensitivity of 78%, and accuracy of 83% and an AUROC of 0.95. As compared to a normal pancreas, it was also used to identify stage I and II PDACs with a sensitivity of 74%, specificity of 75%, and an AUROC of 0.82. In another study, a machine learning algorithm was formulated based on the biochemical differences in the serum of 2 groups of subjects (PDAC group and High risk group) detected *via* the use of Probe Electrospray Ionization Mass Spectrometry (PESI-MS) to identify early stages of pancreatic cancer[62]. It was able to differentiate healthy controls from subjects with earlier stage of PDAC with sensitivity of 81.2% and specificity of 96.8% respectively and an accuracy of 92.9%.

At present, these studies have shown that AI can offer the advantage of identifying specific microRNA and genetic combinations to identifying pancreatic cancer at a faster speed, making this process less laborious. However, these studies lack external validation, limiting their application in modern practice. Besides, studies utilising AI to formulate specific sequences to accurately predict future pancreatic cancer development are still lacking. More studies are required to analyse its ability in predicting future pancreatic cancer for high risk groups especially during the latency period.

Table 1 Studies on artificial intelligence using computed tomography or MRI imaging to diagnose pancreatic ductal adenocarcinoma

Ref.	Clinical question	Training set (number of subjects)	Validation set (number of subjects)	AI instrument	AUROC	Accuracy	Sensitivity	Specificity
Watson *et al* [66], 2021	Detection of pancreatic cystic neoplasms (including PDAC) *vs* benign cysts	18	9	CNN	NA	NA	NA	NA
Si *et al* [65], 2021	Detection of pancreatic cancer (including PDAC, IPMN, PNET)	319	347	DL	0.871	87.6% for PDAC	86.8% for pancreatic cancer	69.5% for pancreatic cancer
Park *et al*[64], 2020	Distinguishing pancreatic cancer tissue from autoimmune pancreatitis	120	62	Random forest machine learning	0.975	95.2%	89.7%	100%
Ma *et al* [63], 2020	Differentiate pancreatic cancer from benign tissue	330	41	CNN	0.9653 (plain scan)	95.47% (plain scan),95.76% (arterial scan), 95.15% (venous phase)	91.58% (plain scan), 94.08% (arterial scan), 92.28% (venous phase)	98.3% (plain scan), 97.6% (arterial scan), 97.9% (venous phase)
Zhang *et al* [67], 2020	Detection of pancreatic cancer	2650 images	240 images	CNN	0.9455	90.2%	83.8%	91.8%
Liu *et al* [69], 2020	Differentiating pancreatic cancer tissue from non-cancerous pancreatic tissue	412	139	CNN	0.92	83.2%	79.0%	97.6%
Gao *et al*[71], 2020	To differentiate pancreatic diseases in pancreatic lesions	398	106	CNN	0.9035 (includes PDAC, adenosquamous carcinoma, acinar cell carcinoma, colloid carcinoma, myoepithelial carcinoma, undifferentiated carcinoma with osteoclast-like giant cells, mucinous cystadenocarcinoma, pancreatoblastoma, pancreatic neuroendocrine carcinoma and metastatic carcinoma)	NA	NA	NA
Chu *et al*[70], 2019	Differentiating PDAC from normal pancreas	255	125	Random forest	NA	93.6%	95%	92.3%
Zhu *et al*[72], 2019	Detecting PDAC from normal pancreas	205	234	CNN	NA	57.3%	94.1%	98.5%
Liu *et al* [73], 2019	Diagnosis of pancreatic cancer	238	100	CNN	0.9632	NA	NA	NA
Corral *et al* [21], 2019	Identify and stratify IPMN lesions	139		DL	0.783	NA	75% (for PDAC or high grade dysplasia)	78% (for PDAC or high grade dysplasia)
Chu *et al*[74], 2019	Differentiating PDAC from normal pancreas	456		DL	NA	NA	94.1%	98.5%
Fu *et al* [75], 2018	Pancreas segmentation (including PDAC, IPMN, Pancreatic Neuroendocrine	59		CNN	NA	NA	82.5%	76.22 (PPV)

	Tumors, Serous Cyst Adenoma, and Solid Pseudopapillary Tumour of the pancreas)				

AUROC: Area under the receiver operating characteristic; AI: Artificial intelligence; CNN: Convolutional neural network; DL: Deep learning; NA: Not available; IPMN: Intraductal papillary mucinous neoplasm; PNET: Pancreatic neuroendocrine tumour; PDAC: Pancreatic ductal adenocarcinoma.

CURRENT EVIDENCE IN PREDICTING THE DEVELOPMENT OF PANCREATIC LESIONS INTO PDAC IN THE FUTURE

Various studies have been conducted using AI to diagnose pancreatic cancer and yielded promising results. Table 1 summarises the studies to date[21,63-75]. In a retrospective study, Liu *et al*[69] was able to train a convolutional neural network (CNN) to identify pancreatic cancer on contrast-enhanced CT and achieve an AUROC of 0.9, with more than 90% for its sensitivity and specificity for its test set. It maintained good sensitivity of 91.3%, specificity of 84.5%, an accuracy of 85.6% and AUROC of 0.955 (95%CI 0.955-0.956) with the validation set. Further analysis revealed that with CNN, radiologists missed 7% of the pancreatic cancers, of which majority were accurately diagnosed by CNN[69]. By enhancing the CNN, Liu *et al*[73] was able to process the CT images and obtain the diagnosis faster than the radiologists (3 s for CNN *vs* 8 mins for a radiologist) with an AUROC of 0.9632, proving that AI is comparable to radiologists.

Besides CT, EUS has been frequently utilised to diagnosed pancreatic cancer. Table 2 summaries these studies[19,20,76-86]. The EUS-CAD based CNN was developed in a retrospective study by Tonozuka *et al*[83] to identify lesions harbouring pancreatic cancer in patients with chronic pancreatitis with a sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of 90.2%, 74.9%, 80.1%, and 88.7%, respectively, and an AUROC of 0.924. Similar findings were also echoed in Zhu *et al*[86] who utilised SVM to obtain a sensitivity, specificity, PPV and NPV of over 90% for diagnosis of pancreatic cancer in chronic pancreatitis.

Despite numerous studies looking at using AI to diagnose pancreatic cancer (as shown in Tables 1 and 2), only a few attempted to predict the development to pancreatic cancer. On average, CT changes for early pancreatic cancer starts approximately 12 to 18 mo before diagnosis[87]. Yet, pancreatic cancer can advance from being undetectable to metastatic in a short period of time even before the next surveillance imaging[88,89]. AI-based imaging itself cannot be used to predict pancreatic cancer and should be combined with other markers.

An ideal AI model for predicting pancreatic cancer is one that integrates multiple biochemical, radiological and clinical data[90]. In a retrospective proof-of-concept study, Springer *et al*[91] developed a supervised machine learning-based approach (CompCyst) based on a combination of patient-reported symptoms, imaging results (including CT, MRI and EUS images), cyst fluid and molecular characteristics to calculate its malignant potential and subsequently determine the management of pancreatic cyst(s). When tested against the validation set, CompCyst outperformed the current standard of care (accuracy 56%) in its ability to identify patients who required surgery, close monitoring or can be discharged (accuracy 69%). CompCyst correctly identified 60% of the surgeries that were not warranted and could have been avoided, while not compromising on its ability to identifying those who truly require surgery. With CompCyst, 71% of the pancreatic lesions were correctly identified as PDAC as compared to 58% based on clinical suspicion[91].

While this study has proven that AI has the potential to incorporate various clinical characteristics, biomarkers, and imaging characteristics to assess for the malignant potential of a pancreatic lesion, it has a number of limitations. Firstly, the imaging characteristics and molecular biomarkers that were identified as high risk features were obtained at the time of surgery and not during screening. These features may not be present early enough to be identified by routine screening. Secondly, important risk factors (including age and diabetes) that were crucial in the early detection of PDAC (as shown in Figure 1) were not included in its learning process, representing a missed step in the screening process. Finally, CompCyst is yet to be externally validated and cannot be applied to the clinical setting currently.

While CompCyst is a potential tool to aid in clinical decision making, future studies aiming at early detection of PDAC face a myriad of challenges. Firstly, the pancreas is a complex organ. Unlike the other organs, the pancreas can be highly variable in its anatomy and location. Moreover, the training data set is highly dependent on the quality of the images provided. Hence, automated segmentation of the pancreas *via* a deep learning approach remains challenging[92]. Secondly, the lack of databases limits the ability to develop new training sets. There are currently only a few open-access databases[93], and there are issues regarding sharing of images across various institutions as pointed out by the Alliance of Pancreatic Cancer Consortia imaging working group[90]. Finally, the algorithm for early detection of PDAC will have to evaluate images of pancreatic lesion(s) across different time points of

Table 2 Studies on artificial intelligence using endoscopic ultrasound to diagnose pancreatic ductal adenocarcinoma

Ref.	Clinical question	Training set (number of subjects)	Validation set (number of subjects)	AI instrument	AUROC	Accuracy	Sensitivity	Specificity
Udristoiu *et al*[84], 2021	Detecting focal pancreatic masses in four EUS imaging modalities	65		CNN and Long Short-term Memory models	0.97	97.6%	98.1%	96.7%
Tonozuka *et al*[83], 2021	Detecting PDAC in patients with normal pancreas/Chronic pancreatitis	92		CNN	0.924	NA	90.2%	74.9%
Marya *et al* [78], 2021	Differentiate AIP from PDAC, chronic pancreatitis and other pancreatic diseases	336	124	CNN	0.976	NA	95%	90%
Kuwahara *et al*[77], 2019	Predicting malignancy in IPMN	50		CNN	0.98	94%	95.7%	92.6%
Ozkan *et al* [80], 2016	Differentiating pancreatic cancer from healthy pancreas	260 images	72 images	ANN	NA	87.5%	83.3%	93.3%
Saftoiu *et al* [81], 2015	Differentiate pancreatic cancer from chronic pancreatitis	117	25	ANN	NA	NA	94.6%	94.4%
Zhu *et al* [86], 2013	Differentiating pancreatic cancer from chronic pancreatitis.	194	194	SVM	NA	94.2%	96.3%	93.4%
Saftoiu *et al* [82], 2012	Diagnosis of focal pancreatic lesions	258 patients		ANN	0.94	84.27%	87.59%	82.94%
Zhang *et al* [85], 2010	Differentiate pancreatic cancer from non-tumorous tissue	108	108	SVM	NA	97.98%	94.3%	99.45%
Saftoiu *et al* [20], 2008 cancer	Differentiate normal pancreas, chronic pancreatitis, pancreatic cancer, and neuroendocrine tumors	68		Neural network	0.847 (for PDAC *vs* chronic pancreatitis)	86.1% (for PDAC *vs* chronic pancreatitis)	93.8% (for PDAC *vs* chronic pancreatitis)	63.6% (for PDAC *vs* chronic pancreatitis)
Das *et al* [19], 2008	Differentiating pancreatic adenocarcinoma from non-neoplastic tissue (includes normal pancreas and chronic pancreatitis)	160	159	ANN	0.93	NA	93%	92%
Norton *et al* [79], 2001	Differentiate malignancy from pancreatitis	35		ML	NA	80%	100%	50%

AUROC: Area under the receiver operating characteristic; AI: Artificial intelligence; CNN: Convolutional neural network; EUS: Endoscopic ultrasound; SVM: Support vector machines; ML: Machine learning; NA: Not available; IPMN: Intraductal papillary mucinous neoplasm; PDAC: Pancreatic ductal adenocarcinoma.

surveillance and from different 3 imaging modalities (namely CT, MRI, and EUS). Unlike CompCyst which looks at images at one time point (*i.e.* at surgery), combining multiple images obtained from periodical surveillance *via* these 3 imaging modalities will require a very large database and multiple layers.

There is a major gap that needs to be bridged before AI systems for early detection of pancreatic cancer can be developed. Given sufficient training data and cooperation, AI-based image analyzers could match or even outperform physicians in image classification and lesion detection[90].

CONCLUSION

Despite the recent advances to predict future PDAC, the use of AI in screening for pancreatic cancer

remains limited in the clinical setting. Much of the efforts are made in the research setting and lack external validation and generalisability. However, this field remains promising as we recognise the challenges ahead to bridge the necessary gaps. The hope to develop an integrated AI model to screen for PDAC remains a reality, and it will play a complementary role in assisting physicians in their clinical decision making process but not replace it.

FOOTNOTES

Author contributions: Lin KW performed the literature search and drafted the manuscript; Ang TL performed the literature search and was involved in the drafting of the manuscript; Li JW conceptualised the project, performed literature search and was involved in the drafting of the manuscript; all authors vetted and approved the final manuscript.

Conflict-of-interest statement: The authors declare that they have no conflict of interest.

Country/Territory of origin: Singapore

ORCID number: Kenneth Weicong Lin 0000-0002-0490-4002; Tiing Leong Ang 0000-0001-9993-8549; James Weiquan Li 0000-0002-5241-4278.

S-Editor: Liu JH
L-Editor: A
P-Editor: Liu JH

REFERENCES

1 **GLOBOCAN**. International Agency for Research on Cancer. 2020
2 **Blackford AL**, Canto MI, Klein AP, Hruban RH, Goggins M. Recent Trends in the Incidence and Survival of Stage 1A Pancreatic Cancer: A Surveillance, Epidemiology, and End Results Analysis. *J Natl Cancer Inst* 2020; **112**: 1162-1169 [PMID: 31958122 DOI: 10.1093/jnci/djaa004]
3 **Huang L**, Jansen L, Balavarca Y, Babaei M, van der Geest L, Lemmens V, Van Eycken L, De Schutter H, Johannesen TB, Primic-Žakelj M, Zadnik V, Besselink MG, Schrotz-King P, Brenner H. Stratified survival of resected and overall pancreatic cancer patients in Europe and the USA in the early twenty-first century: a large, international population-based study. *BMC Med* 2018; **16**: 125 [PMID: 30126408 DOI: 10.1186/s12916-018-1120-9]
4 **Gobbi PG**, Bergonzi M, Comelli M, Villano L, Pozzoli D, Vanoli A, Dionigi P. The prognostic role of time to diagnosis and presenting symptoms in patients with pancreatic cancer. *Cancer Epidemiol* 2013; **37**: 186-190 [PMID: 23369450 DOI: 10.1016/j.canep.2012.12.002]
5 **Ryan DP**, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. *N Engl J Med* 2014; **371**: 1039-1049 [PMID: 25207767 DOI: 10.1056/NEJMra1404198]
6 **Abe T**, Blackford AL, Tamura K, Ford M, McCormick P, Chuidian M, Almario JA, Borges M, Lennon AM, Shin EJ, Klein AP, Hruban RH, Canto MI, Goggins M. Deleterious Germline Mutations Are a Risk Factor for Neoplastic Progression Among High-Risk Individuals Undergoing Pancreatic Surveillance. *J Clin Oncol* 2019; **37**: 1070-1080 [PMID: 30883245 DOI: 10.1200/JCO.18.01512]
7 **Klein AP**. Pancreatic cancer epidemiology: understanding the role of lifestyle and inherited risk factors. *Nat Rev Gastroenterol Hepatol* 2021; **18**: 493-502 [PMID: 34002083 DOI: 10.1038/s41575-021-00457-x]
8 **Capurso G**, Paiella S, Carrara S, Butturini G, Secchettin E, Frulloni L, Zerbi A, Falconi M. Italian registry of families at risk of pancreatic cancer: AISP Familial Pancreatic Cancer Study Group. *Dig Liver Dis* 2020; **52**: 1126-1130 [PMID: 32819857 DOI: 10.1016/j.dld.2020.07.027]
9 **Crippa S**, Bassi C, Salvia R, Malleo G, Marchegiani G, Rebours V, Levy P, Partelli S, Suleiman SL, Banks PA, Ahmed N, Chari ST, Fernández-Del Castillo C, Falconi M. Low progression of intraductal papillary mucinous neoplasms with worrisome features and high-risk stigmata undergoing non-operative management: a mid-term follow-up analysis. *Gut* 2017; **66**: 495-506 [PMID: 26743012 DOI: 10.1136/gutjnl-2015-310162]
10 **Weissman S**, Takakura K, Eibl G, Pandol SJ, Saruta M. The Diverse Involvement of Cigarette Smoking in Pancreatic Cancer Development and Prognosis. *Pancreas* 2020; **49**: 612-620 [PMID: 32433397 DOI: 10.1097/MPA.0000000000001550]
11 **Canto MI**, Almario JA, Schulick RD, Yeo CJ, Klein A, Blackford A, Shin EJ, Sanyal A, Yenokyan G, Lennon AM, Kamel IR, Fishman EK, Wolfgang C, Weiss M, Hruban RH, Goggins M. Risk of Neoplastic Progression in Individuals at High Risk for Pancreatic Cancer Undergoing Long-term Surveillance. *Gastroenterology* 2018; **155**: 740-751.e2 [PMID: 29803839 DOI: 10.1053/j.gastro.2018.05.035]

12 **Aslanian HR**, Lee JH, Canto MI. AGA Clinical Practice Update on Pancreas Cancer Screening in High-Risk Individuals: Expert Review. *Gastroenterology* 2020; **159**: 358-362 [PMID: 32416142 DOI: 10.1053/j.gastro.2020.03.088]

13 **Gonda TA**, Everett JN, Wallace M, Simeone DM; PRECEDE Consortium. Recommendations for a More Organized and Effective Approach to the Early Detection of Pancreatic Cancer From the PRECEDE (Pancreatic Cancer Early Detection) Consortium. *Gastroenterology* 2021; **161**: 1751-1757 [PMID: 34454916 DOI: 10.1053/j.gastro.2021.08.036]

14 **Kenner BJ**, Chari ST, Maitra A, Srivastava S, Cleeter DF, Go VL, Rothschild LJ, Goldberg AE. Early Detection of Pancreatic Cancer-a Defined Future Using Lessons From Other Cancers: A White Paper. *Pancreas* 2016; **45**: 1073-1079 [PMID: 27518362 DOI: 10.1097/MPA.0000000000000701]

15 **Overbeek KA**, Levink IJM, Koopmann BDM, Harinck F, Konings ICAW, Ausems MGEM, Wagner A, Fockens P, van Eijck CH, Groot Koerkamp B, Busch ORC, Besselink MG, Bastiaansen BAJ, van Driel LMJW, Erler NS, Vleggaar FP, Poley JW, Cahen DL, van Hooft JE, Bruno MJ; Dutch Familial Pancreatic Cancer Surveillance Study Group. Long-term yield of pancreatic cancer surveillance in high-risk individuals. *Gut* 2021 [PMID: 33820756 DOI: 10.1016/j.pan.2019.05.303]

16 **Richter AN**, Khoshgoftaar TM. A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. *Artif Intell Med* 2018; **90**: 1-14 [PMID: 30017512 DOI: 10.1016/j.artmed.2018.06.002]

17 **Li JW**, Ang TL. Colonoscopy and artificial intelligence: Bridging the gap or a gap needing to be bridged? *AIGE* 2021; **2**: 36-49 [DOI: 10.37126/aige.v2.i2.36]

18 **Kenner B**, Chari ST, Kelsen D, Klimstra DS, Pandol SJ, Rosenthal M, Rustgi AK, Taylor JA, Yala A, Abul-Husn N, Andersen DK, Bernstein D, Brunak S, Canto MI, Eldar YC, Fishman EK, Fleshman J, Go VLW, Holt JM, Field B, Goldberg A, Hoos W, Iacobuzio-Donahue C, Li D, Lidgard G, Maitra A, Matrisian LM, Poblete S, Rothschild L, Sander C, Schwartz LH, Shalit U, Srivastava S, Wolpin B. Artificial Intelligence and Early Detection of Pancreatic Cancer: 2020 Summative Review. *Pancreas* 2021; **50**: 251-279 [PMID: 33835956 DOI: 10.1097/MPA.0000000000001762]

19 **Das A**, Nguyen CC, Li F, Li B. Digital image analysis of EUS images accurately differentiates pancreatic cancer from chronic pancreatitis and normal tissue. *Gastrointest Endosc* 2008; **67**: 861-867 [PMID: 18179797 DOI: 10.1016/j.gie.2007.08.036]

20 **Săftoiu A**, Vilmann P, Gorunescu F, Gheonea DI, Gorunescu M, Ciurea T, Popescu GL, Iordache A, Hassan H, Iordache S. Neural network analysis of dynamic sequences of EUS elastography used for the differential diagnosis of chronic pancreatitis and pancreatic cancer. *Gastrointest Endosc* 2008; **68**: 1086-1094 [PMID: 18656186 DOI: 10.1016/j.gie.2008.04.031]

21 **Corral JE**, Hussein S, Kandel P, Bolan CW, Bagci U, Wallace MB. Deep Learning to Classify Intraductal Papillary Mucinous Neoplasms Using Magnetic Resonance Imaging. *Pancreas* 2019; **48**: 805-810 [PMID: 31210661 DOI: 10.1097/MPA.0000000000001327]

22 **Momeni-Boroujeni A**, Yousefi E, Somma J. Computer-assisted cytologic diagnosis in pancreatic FNA: An application of neural networks to image analysis. *Cancer Cytopathol* 2017; **125**: 926-933 [PMID: 28885766 DOI: 10.1002/cncy.21915]

23 **The Lustgarten Foundation**. Deep Learning for Radiologists: A Beginner's Guide 2021. Available from: https://www.ctisus.com/responsive/deep-learning/felix.asp

24 **The Lustgarten Foundation**. The Pancreatic Cancer Collective 2021. Available from: https://pancreaticcancercollective.org/

25 **National Institutes of Health**. Early Detection Research Network 2021. Available from: https://edrn.nci.nih.gov/

26 **Nakaura T**, Higaki T, Awai K, Ikeda O, Yamashita Y. A primer for understanding radiology articles about machine learning and deep learning. *Diagn Interv Imaging* 2020; **101**: 765-770 [PMID: 33121910 DOI: 10.1016/j.diii.2020.10.001]

27 **Nakata N**. Recent technical development of artificial intelligence for diagnostic medical imaging. *Jpn J Radiol* 2019; **37**: 103-108 [PMID: 30706381 DOI: 10.1007/s11604-018-0804-6]

28 **Shalev-Shwartz S**, Ben-David S. Understanding machine learning. From theory to algorithms. Understanding Machine Learning: From Theory to Algorithms. 2013 [DOI: 10.1017/cbo9781107298019.022]

29 **van der Sommen F**, de Groof J, Struyvenberg M, van der Putten J, Boers T, Fockens K, Schoon EJ, Curvers W, de With P, Mori Y, Byrne M, Bergman JJGHM. Machine learning in GI endoscopy: practical guidance in how to interpret a novel field. *Gut* 2020; **69**: 2035-2045 [PMID: 32393540 DOI: 10.1136/gutjnl-2019-320466]

30 **Kenner BJ**, Abrams ND, Chari ST, Field BF, Goldberg AE, Hoos WA, Klimstra DS, Rothschild LJ, Srivastava S, Young MR, Go VLW. Early Detection of Pancreatic Cancer: Applying Artificial Intelligence to Electronic Health Records. *Pancreas* 2021; **50**: 916-922 [PMID: 34629446 DOI: 10.1097/MPA.0000000000001882]

31 **Poruk KE**, Firpo MA, Mulvihill SJ. Screening for pancreatic cancer. *Adv Surg* 2014; **48**: 115-136 [PMID: 25293611 DOI: 10.1016/j.yasu.2014.05.004]

32 **US Preventive Services Task Force**, Owens DK, Davidson KW, Krist AH, Barry MJ, Cabana M, Caughey AB, Curry SJ, Doubeni CA, Epling JW Jr, Kubik M, Landefeld CS, Mangione CM, Pbert L, Silverstein M, Simon MA, Tseng CW, Wong JB. Screening for Pancreatic Cancer: US Preventive Services Task Force Reaffirmation Recommendation Statement. *JAMA* 2019; **322**: 438-444 [PMID: 31386141 DOI: 10.1001/jama.2019.10232]

33 **Malhotra A**, Rachet B, Bonaventure A, Pereira SP, Woods LM. Can we screen for pancreatic cancer? *PLoS One* 2021; **16**: e0251876 [PMID: 34077433 DOI: 10.1371/journal.pone.0251876]

34 **Chen Q**, Cherry DR, Nalawade V, Qiao EM, Kumar A, Lowy AM, Simpson DR, Murphy JD. Clinical Data Prediction Model to Identify Patients With Early-Stage Pancreatic Cancer. *JCO Clin Cancer Inform* 2021; **5**: 279-287 [PMID: 33739856 DOI: 10.1200/CCI.20.00137]

35 **Appelbaum L**, Cambronero JP, Stevens JP, Horng S, Pollick K, Silva G, Haneuse S, Piatkowski G, Benhaga N, Duey S, Stevenson MA, Mamon H, Kaplan ID, Rinard MC. Development and validation of a pancreatic cancer risk model for the general population using electronic health records: An observational study. *Eur J Cancer* 2021; **143**: 19-30 [PMID: 33278770 DOI: 10.1016/j.ejca.2020.10.019]

36 **Appelbaum L**, Berg A, Cambronero JP, Dang THY, Jin CC, Zhang L, Kundrot S, Palchuk M, EvansLA, KaplanID, Rinard M. Development of a pancreatic cancer prediction model using a multinational medical records database. *JCO* 2021; **39**: 394 [DOI: 10.1200/jco.2021.39.3_suppl.394]

37 **Muhammad W**, Hart GR, Nartowt B, Farrell JJ, Johung K, Liang Y, Deng J. Pancreatic Cancer Prediction Through an Artificial Neural Network. *Front Artif Intell* 2019; **2**: 2 [PMID: 33733091 DOI: 10.3389/frai.2019.00002]

38 **Placido D**, Yuan B, Hjaltelin JX, Haue AD, Yuan C, Kim J, Umeton R, Antell G, Chowdhury A, Franz A, Brais L, Andrews E, Regev A, Kraft P, WolpinBM, Rosenthal M, Brunak S, Sander C. Pancreatic cancer risk predicted from disease trajectories using deep learning. *BioRxiv* 2021 [DOI: 10.1101/2021.06.27.449937]

39 **Murphy SN**, Weber G, Mendis M, Gainer V, Chueh HC, Churchill S, Kohane I. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc* 2010; **17**: 124-130 [PMID: 20190053 DOI: 10.1136/jamia.2009.000893]

40 **Informatics OHDSA**. Data Standardization 2021. Available from: https://ohdsi.org/data-standardization/

41 **Network TNP-CCR**. Accelerating Data Value Across a National Community Health Center (ADVANCE) Network 2021. Available from: http://advancecollaborative.org/

42 **Locker GY**, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, Somerfield MR, Hayes DF, Bast RC Jr; ASCO. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol* 2006; **24**: 5313-5327 [PMID: 17060676 DOI: 10.1200/JCO.2006.08.2644]

43 **Sekiguchi M**, Matsuda T. Limited usefulness of serum carcinoembryonic antigen and carbohydrate antigen 19-9 levels for gastrointestinal and whole-body cancer screening. *Sci Rep* 2020; **10**: 18202 [PMID: 33097814 DOI: 10.1038/s41598-020-75319-8]

44 **Kriz D**, Ansari D, Andersson R. Potential biomarkers for early detection of pancreatic ductal adenocarcinoma. *Clin Transl Oncol* 2020; **22**: 2170-2174 [PMID: 32447642 DOI: 10.1007/s12094-020-02372-0]

45 **Brezgyte G**, Shah V, Jach D, Crnogorac-Jurcevic T. Non-Invasive Biomarkers for Earlier Detection of Pancreatic Cancer-A Comprehensive Review. *Cancers (Basel)* 2021; **13** [PMID: 34072842 DOI: 10.3390/cancers13112722]

46 **Liao Q**, Zhao YP, Yang YC, Li LJ, Long X, Han SM. Combined detection of serum tumor markers for differential diagnosis of solid lesions located at the pancreatic head. *Hepatobiliary Pancreat Dis Int* 2007; **6**: 641-645 [PMID: 18086633]

47 **Yang Y**, Chen H, Wang D, Luo W, Zhu B, Zhang Z. Diagnosis of pancreatic carcinoma based on combined measurement of multiple serum tumor markers using artificial neural network analysis. *Chin Med J (Engl)* 2014; **127**: 1891-1896 [PMID: 24824251]

48 **Yu J**, Ploner A, Kordes M, Löhr M, Nilsson M, de Maturana MEL, Estudillo L, Renz H, Carrato A, Molero X, Real FX, Malats N, Ye W. Plasma protein biomarkers for early detection of pancreatic ductal adenocarcinoma. *Int J Cancer* 2021; **148**: 2048-2058 [PMID: 33411965 DOI: 10.1002/ijc.33464]

49 **Ray K**. Biomarkers for the early detection of PDAC. *Nat Rev Gastroenterol Hepatol* 2017; **14**: 505

50 **Young MR**, Wagner PD, Ghosh S, Rinaudo JA, Baker SG, Zaret KS, Goggins M, Srivastava S. Validation of Biomarkers for Early Detection of Pancreatic Cancer: Summary of The Alliance of Pancreatic Cancer Consortia for Biomarkers for Early Detection Workshop. *Pancreas* 2018; **47**: 135-141 [PMID: 29346214 DOI: 10.1097/MPA.0000000000000973]

51 **Hasan S**, Jacob R, Manne U, Paluri R. Advances in pancreatic cancer biomarkers. *Oncol Rev* 2019; **13**: 410 [PMID: 31044028 DOI: 10.4081/oncol.2019.410]

52 **Tarasiuk A**, Mackiewicz T, Małecka-Panas E, Fichna J. Biomarkers for early detection of pancreatic cancer - miRNAs as a potential diagnostic and therapeutic tool? *Cancer Biol Ther* 2021; **22**: 347-356 [PMID: 34224317 DOI: 10.1080/15384047.2021.1941584]

53 **Schultz NA**, Dehlendorff C, Jensen BV, Bjerregaard JK, Nielsen KR, Bojesen SE, Calatayud D, Nielsen SE, Yilmaz M, Holländer NH, Andersen KK, Johansen JS. MicroRNA biomarkers in whole blood for detection of pancreatic cancer. *JAMA* 2014; **311**: 392-404 [PMID: 24449318 DOI: 10.1001/jama.2013.284664]

54 **Duell EJ**, Lujan-Barroso L, Sala N, Deitz McElyea S, Overvad K, Tjonneland A, Olsen A, Weiderpass E, Busund LT, Moi L, Muller D, Vineis P, Aune D, Matullo G, Naccarati A, Panico S, Tagliabue G, Tumino R, Palli D, Kaaks R, Katzke VA, Boeing H, Bueno-de-Mesquita HBA, Peeters PH, Trichopoulou A, Lagiou P, Kotanidou A, Travis RC, Wareham N, Khaw KT, Ramon Quiros J, Rodríguez-Barranco M, Dorronsoro M, Chirlaque MD, Ardanaz E, Severi G, Boutron-Ruault MC, Rebours V, Brennan P, Gunter M, Scelo G, Cote G, Sherman S, Korc M. Plasma microRNAs as biomarkers of pancreatic cancer risk in a prospective cohort study. *Int J Cancer* 2017; **141**: 905-915 [PMID: 28542740 DOI: 10.1002/ijc.30790]

55 **Khan IA**, Rashid S, Singh N, Singh V, Gunjan D, Das P, Dash NR, Pandey RM, Chauhan SS, Gupta S, Saraya A. Panel of serum miRNAs as potential non-invasive biomarkers for pancreatic ductal adenocarcinoma. *Sci Rep* 2021; **11**: 2824 [PMID: 33531550 DOI: 10.1038/s41598-021-82266-5]

56 **Shams R**, Saberi S, Zali M, Sadeghi A, Ghafouri-Fard S, Aghdaei HA. Identification of potential microRNA panels for pancreatic cancer diagnosis using microarray datasets and bioinformatics methods. *Sci Rep* 2020; **10**: 7559 [PMID: 32371926 DOI: 10.1038/s41598-020-64569-1]

57 **Cao Z**, Liu C, Xu J, You L, Wang C, Lou W, Sun B, Miao Y, Liu X, Wang X, Zhang T, Zhao Y. Plasma microRNA panels to diagnose pancreatic cancer: Results from a multicenter study. *Oncotarget* 2016; **7**: 41575-41583 [PMID: 27223429 DOI: 10.18632/oncotarget.9491]

58 **Almeida PP**, Cardoso CP, de Freitas LM. PDAC-ANN: an artificial neural network to predict pancreatic ductal adenocarcinoma based on gene expression. *BMC Cancer* 2020; **20**: 82 [PMID: 32005189 DOI: 10.1186/s12885-020-6533-0]

59 **Alizadeh Savareh B**, Asadzadeh Aghdaie H, Behmanesh A, Bashiri A, Sadeghi A, Zali M, Shams R. A machine learning approach identified a diagnostic model for pancreatic cancer through using circulating microRNA signatures. *Pancreatology* 2020; **20**: 1195-1204 [PMID: 32800647 DOI: 10.1016/j.pan.2020.07.399]

60 **Yang J**, Xu R, Wang C, Qiu J, Ren B, You L. Early screening and diagnosis strategies of pancreatic cancer: a comprehensive review. *Cancer Commun (Lond)* 2021; **41**: 1257-1274 [PMID: 34331845 DOI: 10.1002/cac2.12204]

61 **Khatri I**, Bhasin MK. A Transcriptomics-Based Meta-Analysis Combined With Machine Learning Identifies a Secretory Biomarker Panel for Diagnosis of Pancreatic Adenocarcinoma. *Front Genet* 2020; **11**: 572284 [PMID: 33133160 DOI: 10.3389/fgene.2020.572284]

62 **Chung WY**, Correa E, Yoshimura K, Chang MC, Dennison A, Takeda S, Chang YT. Using probe electrospray ionization

mass spectrometry and machine learning for detecting pancreatic cancer with high performance. *Am J Transl Res* 2020; **12**: 171-179 [PMID: 32051746]

63 **Ma H**, Liu ZX, Zhang JJ, Wu FT, Xu CF, Shen Z, Yu CH, Li YM. Construction of a convolutional neural network classifier developed by computed tomography images for pancreatic cancer diagnosis. *World J Gastroenterol* 2020; **26**: 5156-5168 [PMID: 32982116 DOI: 10.3748/wjg.v26.i34.5156]

64 **Park S**, Chu LC, Hruban RH, Vogelstein B, Kinzler KW, Yuille AL, Fouladi DF, Shayesteh S, Ghandili S, Wolfgang CL, Burkhart R, He J, Fishman EK, Kawamoto S. Differentiating autoimmune pancreatitis from pancreatic ductal adenocarcinoma with CT radiomics features. *Diagn Interv Imaging* 2020; **101**: 555-564 [PMID: 32278586 DOI: 10.1016/j.diii.2020.03.002]

65 **Si K**, Xue Y, Yu X, Zhu X, Li Q, Gong W, Liang T, Duan S. Fully end-to-end deep-learning-based diagnosis of pancreatic tumors. *Theranostics* 2021; **11**: 1982-1990 [PMID: 33408793 DOI: 10.7150/thno.52508]

66 **Watson MD**, Lyman WB, Passeri MJ, Murphy KJ, Sarantou JP, Iannitti DA, Martinie JB, Vrochides D, Baker EH. Use of Artificial Intelligence Deep Learning to Determine the Malignant Potential of Pancreatic Cystic Neoplasms With Preoperative Computed Tomography Imaging. *Am Surg* 2021; **87**: 602-607 [PMID: 33131302 DOI: 10.1177/0003134820953779]

67 **Zhang Z**, Li S, Wang Z, Lu Y. A Novel and Efficient Tumor Detection Framework for Pancreatic Cancer *via* CT Images. *Annu Int Conf IEEE Eng Med Biol Soc* 2020; **2020**: 1160-1164 [PMID: 33018193 DOI: 10.1109/EMBC44109.2020.9176172]

68 **Batts KP**, Ludwig J. Chronic hepatitis. An update on terminology and reporting. *Am J Surg Pathol* 1995; **19**: 1409-1417 [PMID: 7503362 DOI: 10.1097/00000478-199512000-00007]

69 **Liu KL**, Wu T, Chen PT, Tsai YM, Roth H, Wu MS, Liao WC, Wang W. Deep learning to distinguish pancreatic cancer tissue from non-cancerous pancreatic tissue: a retrospective study with cross-racial external validation. *Lancet Digit Health* 2020; **2**: e303-e313 [PMID: 33328124 DOI: 10.1016/S2589-7500(20)30078-9]

70 **Chu LC**, Park S, Kawamoto S, Fouladi DF, Shayesteh S, Zinreich ES, Graves JS, Horton KM, Hruban RH, Yuille AL, Kinzler KW, Vogelstein B, Fishman EK. Utility of CT Radiomics Features in Differentiation of Pancreatic Ductal Adenocarcinoma From Normal Pancreatic Tissue. *AJR Am J Roentgenol* 2019; **213**: 349-357 [PMID: 31012758 DOI: 10.2214/AJR.18.20901]

71 **Gao X**, Wang X. Performance of deep learning for differentiating pancreatic diseases on contrast-enhanced magnetic resonance imaging: A preliminary study. *Diagn Interv Imaging* 2020; **101**: 91-100 [PMID: 31375430 DOI: 10.1016/j.diii.2019.07.002]

72 **Zhu Z**, Xia Y, Xie L, Fishman EK, Yuille AL. Multi-Scale Coarse-to-Fine Segmentation for Screening Pancreatic Ductal Adenocarcinoma. MICCAI; 2019

73 **Liu SL**, Li S, Guo YT, Zhou YP, Zhang ZD, Lu Y. Establishment and application of an artificial intelligence diagnosis system for pancreatic cancer with a faster region-based convolutional neural network. *Chin Med J (Engl)* 2019; **132**: 2795-2803 [PMID: 31856050 DOI: 10.1097/CM9.0000000000000544]

74 **Chu LC**, Park S, Kawamoto S, Wang Y, Zhou Y, Shen W, Zhu Z, Xia Y, Xie L, Liu F, Yu Q, Fouladi DF, Shayesteh S, Zinreich E, Graves JS, Horton KM, Yuille AL, Hruban RH, Kinzler KW, Vogelstein B, Fishman EK. Application of Deep Learning to Pancreatic Cancer Detection: Lessons Learned From Our Initial Experience. *J Am Coll Radiol* 2019; **16**: 1338-1342 [PMID: 31492412 DOI: 10.1016/j.jacr.2019.05.034]

75 **Fu M**, Wu W, Hong X, Liu Q, Jiang J, Ou Y, Zhao Y, Gong X. Hierarchical combinatorial deep learning architecture for pancreas segmentation of medical computed tomography cancer images. *BMC Syst Biol* 2018; **12**: 56 [PMID: 29745840 DOI: 10.1186/s12918-018-0572-z]

76 **Angulo P**, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC, Enders F, Saksena S, Burt AD, Bida JP, Lindor K, Sanderson SO, Lenzi M, Adams LA, Kench J, Therneau TM, Day CP. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology* 2007; **45**: 846-854 [PMID: 17393509 DOI: 10.1002/hep.21496]

77 **Kuwahara T**, Hara K, Mizuno N, Okuno N, Matsumoto S, Obata M, Kurita Y, Koda H, Toriyama K, Onishi S, Ishihara M, Tanaka T, Tajika M, Niwa Y. Usefulness of Deep Learning Analysis for the Diagnosis of Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. *Clin Transl Gastroenterol* 2019; **10**: 1-8 [PMID: 31117111 DOI: 10.14309/ctg.0000000000000045]

78 **Marya NB**, Powers PD, Chari ST, Gleeson FC, Leggett CL, Abu Dayyeh BK, Chandrasekhara V, Iyer PG, Majumder S, Pearson RK, Petersen BT, Rajan E, Sawas T, Storm AC, Vege SS, Chen S, Long Z, Hough DM, Mara K, Levy MJ. Utilisation of artificial intelligence for the development of an EUS-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut* 2021; **70**: 1335-1344 [PMID: 33028668 DOI: 10.1136/gutjnl-2020-322821]

79 **Norton ID**, Zheng Y, Wiersema MS, Greenleaf J, Clain JE, Dimagno EP. Neural network analysis of EUS images to differentiate between pancreatic malignancy and pancreatitis. *Gastrointest Endosc* 2001; **54**: 625-629 [PMID: 11677484 DOI: 10.1067/mge.2001.118644]

80 **Ozkan M**, Cakiroglu M, Kocaman O, Kurt M, Yilmaz B, Can G, Korkmaz U, Dandil E, Eksi Z. Age-based computer-aided diagnosis approach for pancreatic cancer on endoscopic ultrasound images. *Endosc Ultrasound* 2016; **5**: 101-107 [PMID: 27080608 DOI: 10.4103/2303-9027.180473]

81 **Săftoiu A**, Vilmann P, Dietrich CF, Iglesias-Garcia J, Hocke M, Seicean A, Ignee A, Hassan H, Streba CT, Ioncică AM, Gheonea DI, Ciurea T. Quantitative contrast-enhanced harmonic EUS in differential diagnosis of focal pancreatic masses (with videos). *Gastrointest Endosc* 2015; **82**: 59-69 [PMID: 25792386 DOI: 10.1016/j.gie.2014.11.040]

82 **Săftoiu A**, Vilmann P, Gorunescu F, Janssen J, Hocke M, Larsen M, Iglesias-Garcia J, Arcidiacono P, Will U, Giovannini M, Dietrich CF, Havre R, Gheorghe C, McKay C, Gheonea DI, Ciurea T; European EUS Elastography Multicentric Study Group. Efficacy of an artificial neural network-based approach to endoscopic ultrasound elastography in diagnosis of focal pancreatic masses. *Clin Gastroenterol Hepatol* 2012; **10**: 84-90.e1 [PMID: 21963957 DOI: 10.1016/j.cgh.2011.09.014]

83 **Tonozuka R**, Itoi T, Nagata N, Kojima H, Sofuni A, Tsuchiya T, Ishii K, Tanaka R, Nagakawa Y, Mukai S. Deep learning analysis for the detection of pancreatic cancer on endosonographic images: a pilot study. *J Hepatobiliary Pancreat Sci*

2021; **28**: 95-104 [PMID: 32910528 DOI: 10.1002/jhbp.825]

84 **Udriștoiu AL**, Cazacu IM, Gruionu LG, Gruionu G, Iacob AV, Burtea DE, Ungureanu BS, Costache MI, Constantin A, Popescu CF, Udriștoiu Ș, Săftoiu A. Real-time computer-aided diagnosis of focal pancreatic masses from endoscopic ultrasound imaging based on a hybrid convolutional and long short-term memory neural network model. *PLoS One* 2021; **16**: e0251701 [PMID: 34181680 DOI: 10.1371/journal.pone.0251701]

85 **Zhang MM**, Yang H, Jin ZD, Yu JG, Cai ZY, Li ZS. Differential diagnosis of pancreatic cancer from normal tissue with digital imaging processing and pattern recognition based on a support vector machine of EUS images. *Gastrointest Endosc* 2010; **72**: 978-985 [PMID: 20855062 DOI: 10.1016/j.gie.2010.06.042]

86 **Zhu M**, Xu C, Yu J, Wu Y, Li C, Zhang M, Jin Z, Li Z. Differentiation of pancreatic cancer and chronic pancreatitis using computer-aided diagnosis of endoscopic ultrasound (EUS) images: a diagnostic test. *PLoS One* 2013; **8**: e63820 [PMID: 23704940 DOI: 10.1371/journal.pone.0063820]

87 **Singh DP**, Sheedy S, Goenka AH, Wells M, Lee NJ, Barlow J, Sharma A, Kandlakunta H, Chandra S, Garg SK, Majumder S, Levy MJ, Takahashi N, Chari ST. Computerized tomography scan in pre-diagnostic pancreatic ductal adenocarcinoma: Stages of progression and potential benefits of early intervention: A retrospective study. *Pancreatology* 2020; **20**: 1495-1501 [PMID: 32950386 DOI: 10.1016/j.pan.2020.07.410]

88 **Yu J**, Blackford AL, Dal Molin M, Wolfgang CL, Goggins M. Time to progression of pancreatic ductal adenocarcinoma from low-to-high tumour stages. *Gut* 2015; **64**: 1783-1789 [PMID: 25636698 DOI: 10.1136/gutjnl-2014-308653]

89 **Overbeek KA**, Goggins MG, Dbouk M, Levink IJM, Koopmann BDM, Chuidian M, Konings ICAW, Paiella S, Earl J, Fockens P, Gress TM, Ausems MGEM, Poley JW, Thosani NC, Half E, Lachter J, Stoffel EM, Kwon RS, Stoita A, Kastrinos F, Lucas AL, Syngal S, Brand RE, Chak A, Carrato A, Vleggaar FP, Bartsch DK, van Hooft JE, Cahen DL, Canto MI, Bruno MJ; International Cancer of the Pancreas Screening Consortium. Timeline of Development of Pancreatic Cancer and Implications for Successful Early Detection in High-Risk Individuals. *Gastroenterology* 2022; **162**: 772-785.e4 [PMID: 34678218 DOI: 10.1053/j.gastro.2021.10.014]

90 **Young MR**, Abrams N, Ghosh S, Rinaudo JAS, Marquez G, Srivastava S. Prediagnostic Image Data, Artificial Intelligence, and Pancreatic Cancer: A Tell-Tale Sign to Early Detection. *Pancreas* 2020; **49**: 882-886 [PMID: 32675784 DOI: 10.1097/MPA.0000000000001603]

91 **Springer S**, Masica DL, Dal Molin M, Douville C, Thoburn CJ, Afsari B, Li L, Cohen JD, Thompson E, Allen PJ, Klimstra DS, Schattner MA, Schmidt CM, Yip-Schneider M, Simpson RE, Fernandez-Del Castillo C, Mino-Kenudson M, Brugge W, Brand RE, Singhi AD, Scarpa A, Lawlor R, Salvia R, Zamboni G, Hong SM, Hwang DW, Jang JY, Kwon W, Swan N, Geoghegan J, Falconi M, Crippa S, Doglioni C, Paulino J, Schulick RD, Edil BH, Park W, Yachida S, Hijioka S, van Hooft J, He J, Weiss MJ, Burkhart R, Makary M, Canto MI, Goggins MG, Ptak J, Dobbyn L, Schaefer J, Sillman N, Popoli M, Klein AP, Tomasetti C, Karchin R, Papadopoulos N, Kinzler KW, Vogelstein B, Wolfgang CL, Hruban RH, Lennon AM. A multimodality test to guide the management of patients with a pancreatic cyst. *Sci Transl Med* 2019; **11** [PMID: 31316009 DOI: 10.1126/scitranslmed.aav4772]

92 **Yang Z**, Zhang L, Zhang M, Feng J, Wu Z, Ren F, Lv Y. Pancreas Segmentation in Abdominal CT Scans using Inter-/Intra-Slice Contextual Information with a Cascade Neural Network. *Annu Int Conf IEEE Eng Med Biol Soc* 2019; **2019**: 5937-5940 [PMID: 31947200 DOI: 10.1109/EMBC.2019.8856774]

93 **Barat M**, Chassagnon G, Dohan A, Gaujoux S, Coriat R, Hoeffel C, Cassinotto C, Soyer P. Artificial intelligence: a critical review of current applications in pancreatic imaging. *Jpn J Radiol* 2021; **39**: 514-523 [PMID: 33550513 DOI: 10.1007/s11604-021-01098-5]

Artificial Intelligence in Medical Imaging

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Med Imaging 2022 April 28; 3(2): 33-41

DOI: 10.35711/aimi.v3.i2.33

ISSN 2644-3260 (online)

MINIREVIEWS

Artificial intelligence: Advances and new frontiers in medical imaging

Marc R Fromherz, Mina S Makary

Specialty type: Radiology, nuclear medicine and medical imaging

Provenance and peer review: Invited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): C, C
Grade D (Fair): 0
Grade E (Poor): 0

P-Reviewer: Dabbakuti JRKKK, India; Liu Y, China; Wan XH, China

Received: December 20, 2021
Peer-review started: December 20, 2021
First decision: January 26, 2022
Revised: February 20, 2022
Accepted: April 21, 2022
Article in press: April 21, 2022
Published online: April 28, 2022

Marc R Fromherz, Department of Radiology, The Ohio State University Wexner Medical Center, Columbus, OH 43210, United States

Mina S Makary, Division of Vascular and Interventional Radiology, Department of Radiology, The Ohio State University Wexner Medical Center, Columbus, OH 43210, United States

Corresponding author: Mina S Makary, MD, Assistant Professor, Attending Doctor, Director, Division of Vascular and Interventional Radiology, Department of Radiology, The Ohio State University Wexner Medical Center, 395 W. 12th Ave 4th Floor Faculty Office Tower, Columbus, OH 43210, United States. mina.makary@osumc.edu

Abstract

Artificial intelligence (AI) has been entwined with the field of radiology ever since digital imaging began replacing films over half a century ago. These algorithms, ranging from simplistic speech-to-text dictation programs to automated interpretation neural networks, have continuously sought to revolutionize medical imaging. With the number of imaging studies outpacing the amount of trained of readers, AI has been implemented to streamline workflow efficiency and provide quantitative, standardized interpretation. AI relies on massive amounts of data for its algorithms to function, and with the wide-spread adoption of Picture Archiving and Communication Systems (PACS), imaging data is accumulating rapidly. Current AI algorithms using machine-learning technology, or computer aided-detection, have been able to successfully pool this data for clinical use, although the scope of these algorithms remains narrow. Many systems have been developed to assist the workflow of the radiologist through PACS optimization and imaging study triage, however interpretation has generally remained a human responsibility for now. In this review article, we will summarize the current successes and limitations of AI in radiology, and explore the exciting prospects that deep-learning technology offers for the future.

Key Words: Artificial intelligence; Machine-learning; Deep-learning; Radiology workflow; Image interpretation

Core Tip: Artificial intelligence (AI) has been an increasingly publicized subject in the field of radiology. This review will attempt to summarize the evolving philosophy and mechanisms behind the AI movement as well as the current applications, limitations, and future directions of the field.

Citation: Fromherz MR, Makary MS. Artificial intelligence: Advances and new frontiers in medical imaging. *Artif Intell Med Imaging* 2022; 3(2): 33-41
URL: https://www.wjgnet.com/2644-3260/full/v3/i2/33.htm
DOI: https://dx.doi.org/10.35711/aimi.v3.i2.33

INTRODUCTION

Advancements in artificial intelligence (AI) technology have created a stir of excitement—and trepidation—amongst professionals in radiology. With the advent of concepts such as machine learning and artificial neural networks promising instantaneous and accurate image interpretation, AI has been heralded as the next step in radiology evolution[1,2]. The ability to reduce image interpretation time and increase detection to levels beyond what is possible for the human eye could create a revolutionary, and increasingly necessary, impact on patient care across all medical disciplines.

AI in radiology has focused on improving three broad principles attributed to human limitations; efficiency, objectivity, and standardization[1,2,3]. Over the past few years there has been a continual increase in imaging orders, and it has been estimated that a radiologist must interpret an image every 3-4 s to match the demand[3,4] This demand, combined with declining reimbursement, has put more pressure on radiologists to increase productivity[5]. Additionally, human and health system variability has long been seen as a potential target to improve standardization across the field. Depending on who the reader is, what hospital system they work for, the time of day, and the number of scans the radiologist has read can result in measurable discrepancies in accuracy and timeliness of image interpretation[3,6,7].

Despite the exciting potential of AI utilization, the fear of algorithms replacing radiologists is ever present. AI companies have grown at an astonishing rate, with 60 new Food and Drug Administration (FDA) approved products in 2020, however the once foreseen AI takeover has not yet manifested[8-10]. Nonetheless, AI is making an impact, just not in the way it was originally planned. A fundamental shift has occurred in recent years in AI implementation, scope, and underlying philosophy. The idea of "replacing radiologists" is not a viable next step in AI evolution, at least for now, and the new philosophy of "working with radiologists" is one that is rapidly gaining traction[11,12]. By examining the current utilizations and limitations of AI in radiology, we can recognize the importance of this fast-rising technology and where the interaction between human and machine may be headed in the future.

CURRENT AI UTILIZATION IN RADIOLOGY

The current state of AI utilization in the field of radiology is variable based on institution, although there are several widely-adopted systems. Aligning with the newer philosophy of "working with radiologists", many of the current AI systems are being used in a limited capacity as tools to enhance the radiologist's workflow. Many of these AI systems fall under the category of "micro-optimizations" [13].

The primary goal for micro-optimization algorithms is to assist the radiologist in his or her daily tasks rather than fully automating the radiologic process. Micro-optimizations can be broken down into two categories; nonpixel-based optimizations and pixel-based optimizations. By using AI to streamline the efficiency and standardization of time-consuming, mundane, or non-interpretive tasks, radiologists can better allocate their time and energy to further focus on image interpretation, consultation, and overall patient care[3,4,14]. Table 1 provides a summary of AI applications for both nonpixel-based and pixel-based optimizations.

Nonpixel-based optimizations

Nonpixel-based optimizations refers to AI assistance in tasks that are not directly related to image interpretation. Some of these tasks include triaging patients, Picture Archiving and Communication Systems (PACS) optimizations, and standardized reporting. As an example, to better triage patients for immediate interpretation AI systems are currently being tested for risk stratification in patients with possible aortic dissection or aneurysm rupture[15,16]. As a different example, through big data analysis, AI algorithms have started to tackle the issue of automated image protocol creations. By reviewing imaging study requests, AI can determine if the study is appropriate, if another study may be more appropriate, or if contrast is necessary or not. With the ability to automatically mine the electronic

Table 1 Areas of radiology workflow with current artificial intelligence implementation

Workflow target	Application examples
Nonpixel-based	
Triage	Risk stratification for aortic pathology and generation of 'aortic calcification score' to assess for disease severity[15,16]
PACS display	Automated hanging protocol and comparison image generation[11]
Order verification	Patient medical record mining with built-in appropriateness criteria guidelines to approve or flag study orders[17,18,19,20,21]
Reporting	Automated data insertion into templates for standardized reporting of chest radiograph findings[23,24]
Pixel-based	
Segmentation	Segmentation of simple lung nodules on chest CT images[43]
Disease registration	PI-RADS lesion classification based on MRI image characteristics[25,26]
Screening	Algorithmic interpretation and classification of screening mammograms[27,28]

PACS: Picture Archiving and Communication Systems; CT: Computed tomography; MRI: magnetic resonance imaging.

medical record system and compare it to established guidelines, the system can then make the appropriate recommendation[17-19]. With an estimated 10% of all imaging studies being ordered in error, these nonpixel-based algorithms can automatically detect and eliminate erroneous study orders [20,21].

The automatic generation of hanging protocols and standardized screen display is another target for optimization. Before data interpretation can commence, a radiologist can spend 10-60 s selecting the appropriate images for comparison[11]. By having the appropriate hanging protocol and display automatically generate, image interpretation can commence instantaneously. What may at first seem like an insignificant amount of time, the elimination of manual protocol selection can significantly improve efficiency and allow for the redirection of the radiologist's brain power toward actual diagnostic interpretation[11].

The standardization of reporting is one of the final areas for optimization, and one that is becoming increasingly necessary among all medical specialties in order to efficiently navigate and report in the electronic medical systems. Reporting is the final step in the radiologist's workflow, and it is also one of the most error-prone[22]. Many micro-optimization AI algorithms are working on increasing the efficiency of reporting through the creation of automatic report generation tools including pre-selected formats specific for the study and automatic annotation. Automating and standardizing reporting can optimize radiologists' reimbursements and save time, as demonstrated by one current chest x-ray reporting algorithm that saved radiologists an average of 8.5 h per month[23,24].

Pixel-based optimizations

While the importance of these nonpixel-based micro-optimizations cannot be understated, the prospect of instantaneous image interpretation is the ultimate ambition of AI. Although AI technology has not yet achieved this ability in a broad sense, the development of pixel-based micro-optimizations have been paramount in maximizing a radiologist's workflow efficiency[14]. Some example applications of these systems involve image segmentation, reconstruction, and disease registration.

AI segmentation has the ability to automatically delineate structures and provide measurements such as organ volume or the surface area of a tumor. Taken a step further, these AI algorithms can be specialized to stage tumors and provide pre-interpreted read-outs such as PI-RADS scores for prostate cancer staging[25,26]. A study by Sanford *et al*[25] demonstrated a modest 40% agreement between an AI algorithm and an expert radiologist when assigning PI-RADS scores based on magnetic resonance imaging (MRI). This result was comparable with previous human inter-reader agreements. Automated segmentation and pre-interpreted read-outs may be maximally utilized in areas that have the most amount of data, such as screening imaging studies.

Utilizing AI for screening processes helps to reduce the workload for radiologists while not over-extending the abilities of AI. As the typical screen produces categorically "positive", "negative", or "inconclusive" results, the complexity of the AI reads can be minimized. Using machine learning for screening detection is referred to as computer aided detection (CADe). CADe is currently being used in screening mammography, where there is an abundance of imaging studies and a relatively disproportionate amount of mammography trained readers[1,2,27]. CADe highlights the area of interest, and it is then determined whether an additional diagnostic study is indicated. CADe for mammography has been around since 1998 and its implementation into clinical workflow has continued to increase allowing radiologists to read more screening studies in less time. Along with the decreased read-time, it should be noted that several studies comparing the accuracy of CADe mammography to traditional radiologist-read mammograms have shown no discernable difference[26]. In one such study, an

ensemble of top-performing AI algorithms combined with a single radiologist reader achieved an area under the curve (AUC) of 0.942, with 92% specificity, outperforming the radiologists' specificity of 90.5%[28]. This is a representative example of new AI algorithms geared toward instantaneous, automatic interpretation.

LIMITATIONS

Despite the constant development of new AI companies, advanced algorithms, and enhanced learning technology, AI has not yet become mainstream in the radiology world due to a combination of both logistical and clinical challenges. The ease of which AI programs can be implemented varies widely based on the scope and technicalities of the clinical problem they aim to solve, as well as the mechanism by which they solve them. In general terminology, there are two main types of AI systems, machine-learning and deep-learning, each of with have some specific limitations of their own[1,29].

Machine-learning AI

Machine-learning functions largely on the principal of pattern recognition. If the machine is able to "see" enough example image characteristics of a certain disease, it can then look at new images and be able to recognize them based on those previously defined features. The caveat here, is that these "pre-defined features", such as tumor volume, density, *etc*., must be hand-fed into each specific machine-learning classifier[3]. In this way the AI does not actually learn, but rather applies the specifics of its pre-engineered programming. Consequently, machine-learning AI is intrinsically limited by these specific characteristics which can reduce its ability to recognize image features, such as rare or unusual disease presentations[30,31]. Figure 1 demonstrates a schematic example of how machine-learning AI systems utilize these pre-defined features for classification. Furthermore, as the breadth of medical knowledge continues to expand, previous CAD systems may become outdated, and therefore obsolete[30]. The theoretical solution to these hard-wired restrictions is the use of AI algorithms that do not rely on pre-engineered feature recognition, but rather one that can learn and adapt in a manner similar to the human brain.

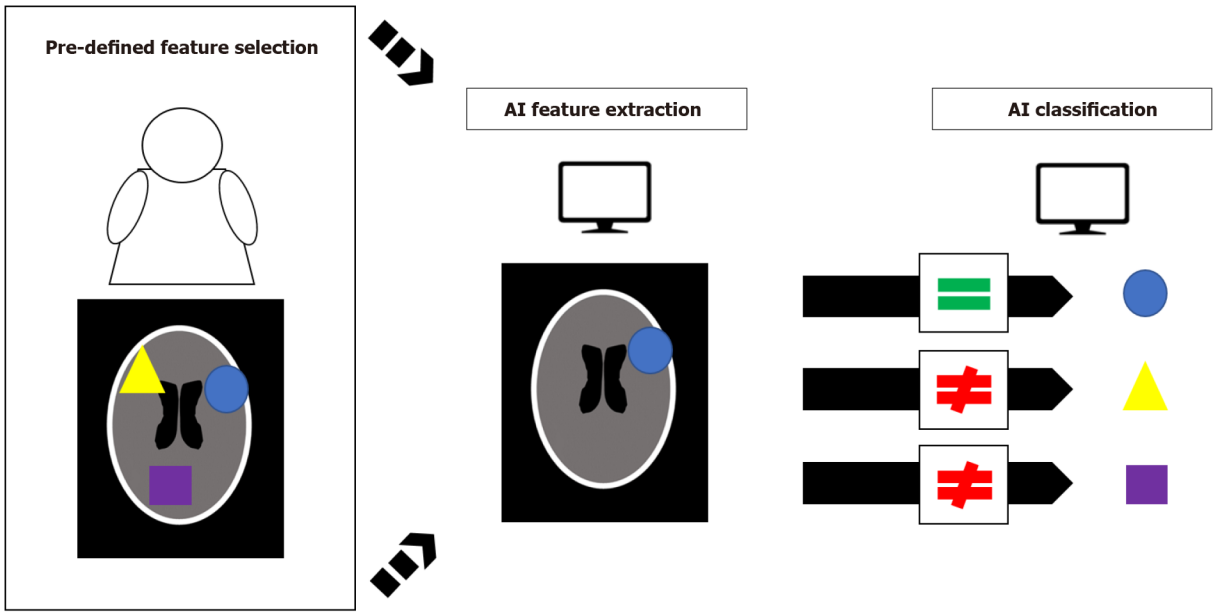
Deep-learning AI

Deep-learning is programmed to mimic the pattern of neural networks such as those in the human brain, referred to in the literature as convolutional neural networks (CNNs). The principal mechanism behind AI algorithms relies on a vast quantity of data, and through this data the AI can develop its own pattern of feature recognition without the need for pre-programming from human experts. Deep-learning AI uses these features to create connections and draw conclusions in a way similar to the human brain, and allowing it to operate freely from human input thus increasing its automaticity and decreasing restrictions[3,32,33]. While in theory this method appears to be a step-up from classical machine-learning technology, the reliance on data and complexity of the mechanism has its limitations.

With the wide-implementation of PACS and an ever-increasing number of medical images, there is no shortage of data for AI algorithms to mine[34]. The issue is not quantity—but quality. Different PACS, different imaging machine manufacturers, and different protocols can all effect the generalizability of an AI algorithm. These variations in image reconstruction, segmentation, and labelling can have adverse effects on the AI's ability to learn, and the process of standardization across these variables would be a time-consuming and expensive task. This is one of the reasons for the current narrow use of AI in clinical practice. Currently approved AI programs only function with specific computed tomography (CT) imager models, specific PAC systems, and specific disease processes. With such a narrow clinical window, AI in its current form is limited in scope[30,31]. If multiple different AI systems are needed for each specific pathology the process of creating and implementing these systems may not be fiscally feasible[35]. Even with implementation, a lapse in the detection of rare diseases would still exist.

Industry acceptance

Questions regarding the mechanism of how deep-learning functions can also create additional limitations, specifically regarding FDA approval and the accuracy of the AI's results[8,36]. The mechanism is extremely complex, and in many instances, the exact way in which the AI forms these CNNs is either unknown or proprietary. If the way the AI algorithm functions to produce its results is not well understood, this begs the question of whether or not its results can be trusted[8,36,37]. This question has haunted AI since its inception, and the answer of whether or not health professionals and patients would be willing to put their faith in the recommendation of a 100% computer-controlled radiologic study is not an easy one to answer. A variety of comparison studies have been conducted to determine whether AI accuracy is comparable to that of human readers, and the results have been mixed.



DOI: 10.35711/aimi.v3.i2.33

Figure 1 Machine-learning requires pre-defined feature inputs which are then extracted in order to classify target image characteristics.
AI: Artificial intelligence.

In the previously mentioned Schaffter *et al*[28] study on breast cancer detection, no single AI algorithm was able to outperform the radiologists, with a specificity of 66.1% for the top-performing algorithm compared to 90.5% for the radiologists. In a breast cancer detection study using a different AI system, the AI outperformed the radiologists with an AUC of 0.740 compared to the radiologists' AUC of 0.625[38]. In a study comparing chest radiograph interpretation, AI outperformed the radiologists on detection of pulmonary edema, underperformed on detection of consolidation, and had comparable performance for detection of pleural effusions[39]. These studies collectively demonstrate that AI systems have mixed performance compared to human radiologists.

The utilization of different algorithms, training datasets, and radiologist experience in these studies makes drawing conclusions about AI's general trustworthiness difficult. Concerns such as these are why the shift toward micro-optimizations has been an attractive one for the interim, however as new technologies are developed and deep-learning systems are polished the future of AI continues to push the boundaries of possibility.

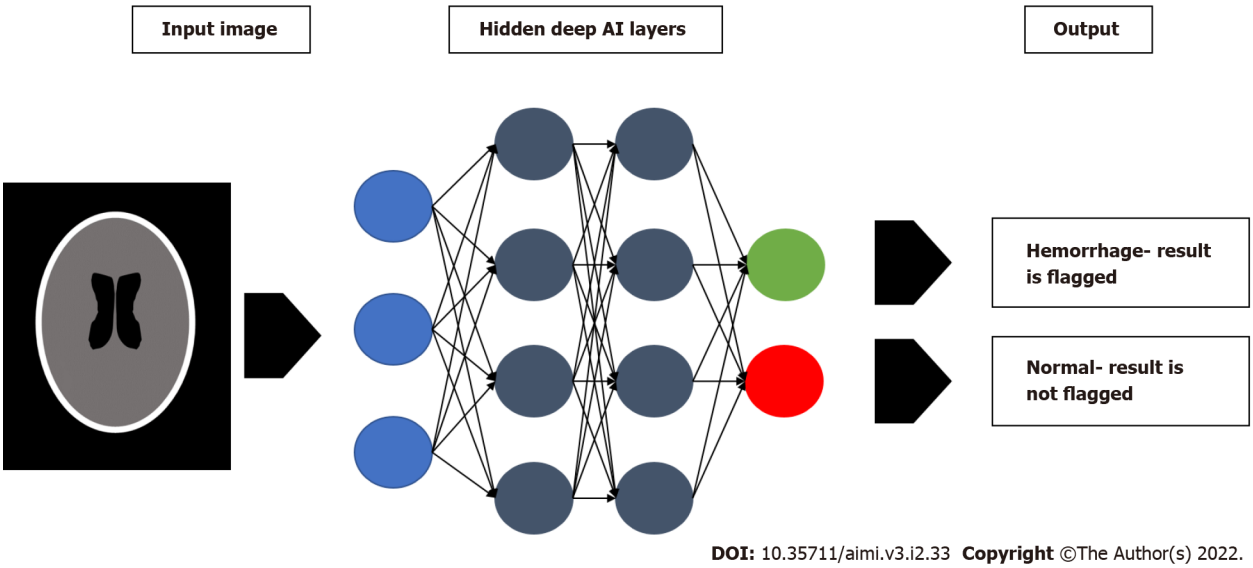
FUTURE DIRECTIONS

The future of AI in radiology is constantly evolving, and with new computer systems, implementation targets, and algorithms being developed seemingly by the day there is no discernable end to what is possible[8-10]. Within PACS, the utilization of deep learning AI could theoretically be implemented wherever large quantities of data are available, although as previously stated there are several limitations to deep learning technology. With the interconnectivity, digitization, and increasing data pool in modern radiology, the limitations of deep-learning may slowly start to be overcome, and the use of micro-optimization may ramp up in scale.

The next phase in AI utilization will likely continue the trend of micro-optimization, but with increased efficiency. As hospital systems become more integrated, with imaging devices and PACS being able to directly communicate with each other, it would only make sense that the AI algorithms within these systems do the same. With AI's current narrow clinical usage, each system excels at only one specific task[30,31]. By combining these systems, the scope of each can be summated into a larger, more efficient system. For example a lung cancer screening CT reconstruction algorithm could be used alongside a hanging protocol algorithm, with CADe for detection, and another algorithm for report generation[40]. Until a more encompassing system is created, combining existing micro-optimizations can scale efficiency in clinical workflow.

Disease recognition and triage

Despite the profound promise of deep learning, it has yet to have seen wide-spread clinical utilization. That being said, the power behind deep learning is data and the amount of available data is continuously growing. As we gather more high-quality data, the deep learning systems should become



DOI: 10.35711/aimi.v3.i2.33

Figure 2 Deep-learning artificial intelligence application in triaging head computed tomography images. The input image characteristics are extracted and analyzed by the convolutional neural network to create an output. The output is then flagged or not flagged depending on the algorithm's interpretation. AI: Artificial intelligence.

more powerful, increasing their usage potential. The full potential of deep learning is still unknown, however there are several promising applications in detection and automated disease monitoring. One of these applications is in the identification of incidental findings. When a radiologist is examining a trauma study, the AI system can detect incidental pulmonary nodules, allowing the radiologist to focus on the primary clinical issue without overlooking other findings[41,42,43]. Looking to improve upon current CAD systems, utilizing deep learning AI for triage is another attractive target, where the urgency of a given study is prioritized and then sent to a radiologist for final interpretation. These algorithms pool hundreds of thousands of imaging studies along with their subsequent reports, and use this information to train their CNNs. In a study of one such algorithm on assigning priority to adult chest radiographs, AI was able to assign priority with a sensitivity of 71% and a specificity of 95%. Importantly, the time taken to report critical findings was reduced significantly from 11.2 to 2.7[32]. Another study on triaging patients based on head CT findings produced similar results, with an AUC of 0.92 for accurately detecting intracranial hemorrhage[44]. Figure 2 is schematic example demonstrating this type of AI triage system. The ability for the system to distinguish between 'normal' and 'abnormal' accurately, and then further stratify 'abnormal' into severity categories, is a promising step toward automated interpretation[32,44].

Disease monitoring

The prospect of monitoring disease progression is a more complicated one, but the ability of the deep learning system to accumulate and track data changes over time makes this an attractive target. These systems may also have the ability to automatically adjust for changes in patient position or body habitus at the times the studies were conducted[3]. One of the obvious applications for this is oncology, with AI models already demonstrating their ability to accurately measure therapeutic response and tumor recurrence[45,46]. Throughout the coronavirus disease 2019 (COVID-19) pandemic, the ability to track disease progression has been crucial for medical decision making. Unfortunately, the wide variability in an individual's disease course has been difficult to predict. To solve this problem, several deep learning systems have been tested to identify minute chest CT changes based on quantitative pixel analysis, giving us a more sophisticated look into the pathophysiology of the disease[47-49]. Not only does this present the potential to make educated decisions for COVID-19 patients regarding the need for hospitalization and allocation of resources, but the pandemic in general has further stressed the need of increased efficiency in radiology during times of unprecedented volume.

CONCLUSION

As the role of AI in radiology continues to advance and diversify, the potential for revolutionary clinical impact persists. One of the most important factors for the continued development of AI in radiology is achieving wide-spread implementation, and to achieve this AI must be embraced by radiologists. Currently, only an estimated 30% of radiologists use AI in day-to-day workflow[50]. With the shift of AI philosophy away from replacing radiologists, the view of AI as a threat to fear may be replaced with its

view as a tool to exploit. As more algorithms are approved, more studies published, and more systems implemented into clinical practice, radiologists and trainees alike need to educate themselves on what AI can do for them and their patients. When radiologists and AI learn to work together, the potential clinical benefits of a human-machine symbiosis can be fully realized.

FOOTNOTES

Author contributions: All authors contributed to the writing and preparation of the manuscript.

Conflict-of-interest statement: The authors declare they have no conflicts of interest.

Country/Territory of origin: United States

ORCID number: Marc R Fromherz 0000-0002-0458-0749; Mina S Makary 0000-0002-2498-7132.

S-Editor: Liu JH
L-Editor: A
P-Editor: Liu JH

REFERENCES

1 **Oakden-Rayner L**. The Rebirth of CAD: How Is Modern AI Different from the CAD We Know? *Radiol Artif Intell* 2019; **1**: e180089 [PMID: 33937793 DOI: 10.1148/ryai.2019180089]
2 **Mun SK**, Wong KH, Lo SB, Li Y, Bayarsaikhan S. Artificial Intelligence for the Future Radiology Diagnostic Service. *Front Mol Biosci* 2020; **7**: 614258 [PMID: 33585563 DOI: 10.3389/fmolb.2020.614258]
3 **Hosny A**, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer* 2018; **18**: 500-510 [PMID: 29777175 DOI: 10.1038/s41568-018-0016-5]
4 **Zha N**, Patlas MN, Duszak R Jr. Radiologist Burnout Is Not Just Isolated to the United States: Perspectives From Canada. *J Am Coll Radiol* 2019; **16**: 121-123 [PMID: 30236858 DOI: 10.1016/j.jacr.2018.07.010]
5 **Stempniak M**. "Grave Concern": Radiologist Reimbursement Expected to Plummet after CMS Clinical Labor Wage Update, August 16, 2021. Available from: https://www.radiologybusiness.com/topics/economics/radiologist-reimbursement-cms-clinical-labor-wage
6 **Patel AG**, Pizzitola VJ, Johnson CD, Zhang N, Patel MD. Radiologists Make More Errors Interpreting Off-Hours Body CT Studies during Overnight Assignments as Compared with Daytime Assignments. *Radiology* 2020; **297**: E281 [PMID: 33074783 DOI: 10.1148/radiol.2020209018]
7 **Hanna TN**, Lamoureux C, Krupinski EA, Weber S, Johnson JO. Effect of Shift, Schedule, and Volume on Interpretive Accuracy: A Retrospective Analysis of 2.9 Million Radiologic Examinations. *Radiology* 2018; **287**: 205-212 [PMID: 29156150 DOI: 10.1148/radiol.2017170555]
8 **Benjamens S**, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med* 2020; **3**: 118 [PMID: 32984550 DOI: 10.1038/s41746-020-00324-0]
9 **Chen MM**, Golding LP, Nicola GN. Who Will Pay for AI? *Radiol Artif Intell* 2021; **3**: e210030 [PMID: 34142090 DOI: 10.1148/ryai.2021210030]
10 **Bloom J**, Dyrda L. 100+ artificial intelligence companies to know in healthcare. Beckers Hospital Review. Published July 19, 2019. Accessed August 4, 2020. Available from: https://www.beckershospitalreview.com/Lists/100-artificial-intelligence-companies-to-know-in-healthcare-2019.html
11 **Langlotz CP**. Will Artificial Intelligence Replace Radiologists? *Radiol Artif Intell* 2019; **1**: e190058 [PMID: 33937794 DOI: 10.1148/ryai.2019190058]
12 **Siwicki B**. Mass General Brigham and the Future of AI in Radiology. Healthcare IT News, May 10, 2021. Available from: https://www.healthcareitnews.com/news/mass-general-brigham-and-future-ai-radiology
13 **Gimenez F**. Real-Time Interpretation: The next Frontier in Radiology AI - MedCity News. Medcitynews, July 22, 2022. Available from: https://medcitynews.com/2021/07/real-time-interpretation-the-next-frontier-in-radiology-ai/
14 **Do HM**, Spear LG, Nikpanah M, Mirmomen SM, Machado LB, Toscano AP, Turkbey B, Bagheri MH, Gulley JL, Folio LR. Augmented Radiologist Workflow Improves Report Value and Saves Time: A Potential Model for Implementation of Artificial Intelligence. *Acad Radiol* 2020; **27**: 96-105 [PMID: 31818390 DOI: 10.1016/j.acra.2019.09.014]
15 **Hahn LD**, Baeumler K, Hsiao A. Artificial intelligence and machine learning in aortic disease. *Curr Opin Cardiol* 2021; **36**: 695-703 [PMID: 34369401 DOI: 10.1097/HCO.0000000000000903]
16 **Chin CW**, Pawade TA, Newby DE, Dweck MR. Risk Stratification in Patients With Aortic Stenosis Using Novel Imaging Approaches. Circ Cardiovasc Imaging 2015; 8: e003421 [PMID: 26198161 DOI: 10.1161/CIRCIMAGING.115.003421]

17 **Desai V**, Flanders A, Zoga AC. Leveraging Technology to Improve Radiology Workflow. *Semin Musculoskelet Radiol* 2018; **22**: 528-539 [PMID: 30399617 DOI: 10.1055/s-0038-1673385]

18 **Hassanpour S**, Langlotz CP. Information extraction from multi-institutional radiology reports. *Artif Intell Med* 2016; **66**: 29-39 [PMID: 26481140 DOI: 10.1016/j.artmed.2015.09.007]

19 **Bhatia N**, Trivedi H, Safdar N, Heilbrun ME. Artificial Intelligence in Quality Improvement: Reviewing Uses of Artificial Intelligence in Noninterpretative Processes from Clinical Decision Support to Education and Feedback. *J Am Coll Radiol* 2020; **17**: 1382-1387 [PMID: 33153542 DOI: 10.1016/j.jacr.2020.08.002]

20 **Bernardy M**, Ullrich CG, Rawson JV, Allen B Jr, Thrall JH, Keysor KJ, James C, Boyes JA, Saunders WM, Lomers W, Mollura DJ, Pyatt RS Jr, Taxin RN, Mabry MR. Strategies for managing imaging utilization. *J Am Coll Radiol* 2009; **6**: 844-850 [PMID: 19945039 DOI: 10.1016/j.jacr.2009.08.003]

21 **Lehnert BE**, Bree RL. Analysis of appropriateness of outpatient CT and MRI referred from primary care clinics at an academic medical center: how critical is the need for improved decision support? *J Am Coll Radiol* 2010; **7**: 192-197 [PMID: 20193924 DOI: 10.1016/j.jacr.2009.11.010]

22 **Onder O**, Yarasir Y, Azizova A, Durhan G, Onur MR, Ariyurek OM. Errors, discrepancies and underlying bias in radiology with case examples: a pictorial review. *Insights Imaging* 2021; **12**: 51 [PMID: 33877458 DOI: 10.1186/s13244-021-00986-8]

23 **Lakhani P**, Prater AB, Hutson RK, Andriole KP, Dreyer KJ, Morey J, Prevedello LM, Clark TJ, Geis JR, Itri JN, Hawkins CM. Machine Learning in Radiology: Applications Beyond Image Interpretation. *J Am Coll Radiol* 2018; **15**: 350-359 [PMID: 29158061 DOI: 10.1016/j.jacr.2017.09.044]

24 **Chung CY**, Makeeva V, Yan J, Prater AB, Duszak R Jr, Safdar NM, Heilbrun ME. Improving Billing Accuracy Through Enterprise-Wide Standardized Structured Reporting With Cross-Divisional Shared Templates. *J Am Coll Radiol* 2020; **17**: 157-164 [PMID: 31918874 DOI: 10.1016/j.jacr.2019.08.034]

25 **Sanford T**, Harmon SA, Turkbey EB, Kesani D, Tuncer S, Madariaga M, Yang C, Sackett J, Mehralivand S, Yan P, Xu S, Wood BJ, Merino MJ, Pinto PA, Choyke PL, Turkbey B. Deep-Learning-Based Artificial Intelligence for PI-RADS Classification to Assist Multiparametric Prostate MRI Interpretation: A Development Study. *J Magn Reson Imaging* 2020; **52**: 1499-1507 [PMID: 32478955 DOI: 10.1002/jmri.27204]

26 **Winkel DJ**, Wetterauer C, Matthias MO, Lou B, Shi B, Kamen A, Comaniciu D, Seifert HH, Rentsch CA, Boll DT. Autonomous Detection and Classification of PI-RADS Lesions in an MRI Screening Population Incorporating Multicenter-Labeled Deep Learning and Biparametric Imaging: Proof of Concept. Diagnostics (Basel) 2020; 10 [PMID: 33202680 DOI: 10.3390/diagnostics10110951]

27 **Raya-Povedano JL**, Romero-Martín S, Elías-Cabot E, Gubern-Mérida A, Rodríguez-Ruiz A, Álvarez-Benito M. AI-based Strategies to Reduce Workload in Breast Cancer Screening with Mammography and Tomosynthesis: A Retrospective Evaluation. *Radiology* 2021; **300**: 57-65 [PMID: 33944627 DOI: 10.1148/radiol.2021203555]

28 **Schaffter T**, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, Lotter W, Jie Z, Du H, Wang S, Feng J, Feng M, Kim HE, Albiol F, Albiol A, Morrell S, Wojna Z, Ahsen ME, Asif U, Jimeno Yepes A, Yohanandan S, Rabinovici-Cohen S, Yi D, Hoff B, Yu T, Chaibub Neto E, Rubin DL, Lindholm P, Margolies LR, McBride RB, Rothstein JH, Sieh W, Ben-Ari R, Harrer S, Trister A, Friend S, Norman T, Sahiner B, Strand F, Guinney J, Stolovitzky G; and the DM DREAM Consortium, Mackey L, Cahoon J, Shen L, Sohn JH, Trivedi H, Shen Y, Buturovic L, Pereira JC, Cardoso JS, Castro E, Kalleberg KT, Pelka O, Nedjar I, Geras KJ, Nensa F, Goan E, Koitka S, Caballero L, Cox DD, Krishnaswamy P, Pandey G, Friedrich CM, Perrin D, Fookes C, Shi B, Cardoso Negrie G, Kawczynski M, Cho K, Khoo CS, Lo JY, Sorensen AG, Jung H. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. *JAMA Netw Open* 2020; **3**: e200265 [PMID: 32119094 DOI: 10.1001/jamanetworkopen.2020.0265]

29 **Do S**, Song KD, Chung JW. Basics of Deep Learning: A Radiologist's Guide to Understanding Published Radiology Articles on Deep Learning. *Korean J Radiol* 2020; **21**: 33-41 [PMID: 31920027 DOI: 10.3348/kjr.2019.0312]

30 **Futoma J**, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. *Lancet Digit Health* 2020; **2**: e489-e492 [PMID: 32864600 DOI: 10.1016/S2589-7500(20)30186-2]

31 **Zech JR**, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS Med* 2018; **15**: e1002683 [PMID: 30399157 DOI: 10.1371/journal.pmed.1002683]

32 **Annarumma M**, Withey SJ, Bakewell RJ, Pesce E, Goh V, Montana G. Automated Triaging of Adult Chest Radiographs with Deep Artificial Neural Networks. *Radiology* 2019; **291**: 196-202 [PMID: 30667333 DOI: 10.1148/radiol.2018180921]

33 **Shen D**, Wu G, Suk HI. Deep Learning in Medical Image Analysis. *Annu Rev Biomed Eng* 2017; **19**: 221-248 [PMID: 28301734 DOI: 10.1146/annurev-bioeng-071516-044442]

34 **Barlow RD**. Larger Volume Data Sets Redefining PACS. Imaging Technology News, June 12, 2008. Available from: http://www.itnonline.com/article/Larger-volume-data-sets-redefining-pacs

35 **Tadavarthi Y**, Vey B, Krupinski E, Prater A, Gichoya J, Safdar N, Trivedi H. The State of Radiology AI: Considerations for Purchase Decisions and Current Market Offerings. *Radiol Artif Intell* 2020; **2**: e200004 [PMID: 33937846 DOI: 10.1148/ryai.2020200004]

36 **Aggarwal R**, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, Ashrafian H, Darzi A. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ Digit Med* 2021; **4**: 65 [PMID: 33828217 DOI: 10.1038/s41746-021-00438-z]

37 **Chartrand G**, Cheng PM, Vorontsov E, Drozdzal M, Turcotte S, Pal CJ, Kadoury S, Tang A. Deep Learning: A Primer for Radiologists. *Radiographics* 2017; **37**: 2113-2131 [PMID: 29131760 DOI: 10.1148/rg.2017170077]

38 **McKinney SM**, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesus M, Corrado GS, Darzi A, Etemadi M, Garcia-Vicente F, Gilbert FJ, Halling-Brown M, Hassabis D, Jansen S, Karthikesalingam A, Kelly CJ, King D, Ledsam JR, Melnick D, Mostofi H, Peng L, Reicher JJ, Romera-Paredes B, Sidebottom R, Suleyman M, Tse D, Young KC, De Fauw J, Shetty S. International evaluation of an AI system for breast cancer screening. *Nature* 2020; **577**: 89-94 [PMID: 31894144 DOI: 10.1038/s41586-019-1799-6]

39 **Wu JT**, Wong KCL, Gur Y, Ansari N, Karargyris A, Sharma A, Morris M, Saboury B, Ahmad H, Boyko O, Syed A,

Jadhav A, Wang H, Pillai A, Kashyap S, Moradi M, Syeda-Mahmood T. Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents. *JAMA Netw Open* 2020; **3**: e2022779 [PMID: 33034642 DOI: 10.1001/jamanetworkopen.2020.22779]

40 **Svoboda E**. Artificial intelligence is improving the detection of lung cancer. *Nature* 2020; **587**: S20-S22 [PMID: 33208974 DOI: 10.1038/d41586-020-03157-9]

41 **Cui S**, Ming S, Lin Y, Chen F, Shen Q, Li H, Chen G, Gong X, Wang H. Development and clinical application of deep learning model for lung nodules screening on CT images. *Sci Rep* 2020; **10**: 13657 [PMID: 32788705 DOI: 10.1038/s41598-020-70629-3]

42 **Ottawa (ON): Canadian Agency for Drugs and Technologies in Health**. Artificial Intelligence for Classification of Lung Nodules: A Review of Clinical Utility, Diagnostic Accuracy, Cost-Effectiveness, and Guidelines [Internet]. 2020-Jan-22 [PMID: 33074628]

43 **Weikert T**, Akinci D'Antonoli T, Bremerich J, Stieltjes B, Sommer G, Sauter AW. Evaluation of an AI-Powered Lung Nodule Algorithm for Detection and 3D Segmentation of Primary Lung Tumors. *Contrast Media Mol Imaging* 2019; **2019**: 1545747 [PMID: 31354393 DOI: 10.1155/2019/1545747]

44 **Chilamkurthy S**, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, Mahajan V, Rao P, Warier P. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet* 2018; **392**: 2388-2396 [PMID: 30318264 DOI: 10.1016/S0140-6736(18)31645-3]

45 **Liu Z**, Wang S, Dong D, Wei J, Fang C, Zhou X, Sun K, Li L, Li B, Wang M, Tian J. The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. *Theranostics* 2019; **9**: 1303-1322 [PMID: 30867832 DOI: 10.7150/thno.30309]

46 **Bi WL**, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, Allison T, Arnaout O, Abbosh C, Dunn IF, Mak RH, Tamimi RM, Tempany CM, Swanton C, Hoffmann U, Schwartz LH, Gillies RJ, Huang RY, Aerts HJWL. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin* 2019; **69**: 127-157 [PMID: 30720861 DOI: 10.3322/caac.21552]

47 **Wynants L**, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, Bonten MMJ, Dahly DL, Damen JAA, Debray TPA, de Jong VMT, De Vos M, Dhiman P, Haller MC, Harhay MO, Henckaerts L, Heus P, Kammer M, Kreuzberger N, Lohmann A, Luijken K, Ma J, Martin GP, McLernon DJ, Andaur Navarro CL, Reitsma JB, Sergeant JC, Shi C, Skoetz N, Smits LJM, Snell KIE, Sperrin M, Spijker R, Steyerberg EW, Takada T, Tzoulaki I, van Kuijk SMJ, van Bussel B, van der Horst ICC, van Royen FS, Verbakel JY, Wallisch C, Wilkinson J, Wolff R, Hooft L, Moons KGM, van Smeden M. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020; **369**: m1328 [PMID: 32265220 DOI: 10.1136/bmj.m1328]

48 **Huang L**, Han R, Ai T, Yu P, Kang H, Tao Q, Xia L. Serial Quantitative Chest CT Assessment of COVID-19: A Deep Learning Approach. *Radiol Cardiothorac Imaging* 2020; **2**: e200075 [PMID: 33778562 DOI: 10.1148/ryct.2020200075]

49 **Li Z**, Zhong Z, Li Y, Zhang T, Gao L, Jin D, Sun Y, Ye X, Yu L, Hu Z, Xiao J, Huang L, Tang Y. From community-acquired pneumonia to COVID-19: a deep learning-based method for quantitative analysis of COVID-19 on thick-section CT scans. *Eur Radiol* 2020; **30**: 6828-6837 [PMID: 32683550 DOI: 10.1007/s00330-020-07042-x]

50 **Allen B**, Agarwal S, Coombs L, Wald C, Dreyer K. 2020 ACR Data Science Institute Artificial Intelligence Survey. *J Am Coll Radiol* 2021; **18**: 1153-1159 [PMID: 33891859 DOI: 10.1016/j.jacr.2021.04.002]

Artificial Intelligence in Medical Imaging

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Med Imaging 2022 April 28; 3(2): 42-54

DOI: 10.35711/aimi.v3.i2.42

ISSN 2644-3260 (online)

SYSTEMATIC REVIEWS

Applications of artificial intelligence in lung ultrasound: Review of deep learning methods for COVID-19 fighting

Laura De Rosa, Serena L'Abbate, Claudia Kusmic, Francesco Faita

Specialty type: Radiology, nuclear medicine and medical imaging

Provenance and peer review: Invited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): A
Grade B (Very good): B
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

P-Reviewer: Haurylenka D, Belarus; Zhang W, China

Received: December 19, 2021
Peer-review started: December 19, 2021
First decision: February 10, 2022
Revised: February 22, 2022
Accepted: April 27, 2022
Article in press: April 27, 2022
Published online: April 28, 2022

Laura De Rosa, Serena L'Abbate, Claudia Kusmic, Francesco Faita, Institute of Clinical Physiology, Consiglio Nazionale delle Ricerche, Pisa 56124, Italy

Serena L'Abbate, Institute of Life Sciences, Scuola Superiore Sant'Anna, Pisa 56124, Italy

Corresponding author: Claudia Kusmic, MSc, PhD, Research Scientist, Institute of Clinical Physiology, Consiglio Nazionale delle Ricerche, Via Giuseppe Moruzzi 1, Pisa 56124, Italy. kusmic@ifc.cnr.it

Abstract

BACKGROUND

The pandemic outbreak of the novel coronavirus disease (COVID-19) has highlighted the need to combine rapid, non-invasive and widely accessible techniques with the least risk of patient's cross-infection to achieve a successful early detection and surveillance of the disease. In this regard, the lung ultrasound (LUS) technique has been proved invaluable in both the differential diagnosis and the follow-up of COVID-19 patients, and its potential may be destined to evolve. Recently, indeed, LUS has been empowered through the development of automated image processing techniques.

AIM

To provide a systematic review of the application of artificial intelligence (AI) technology in medical LUS analysis of COVID-19 patients using the preferred reporting items of systematic reviews and meta-analysis (PRISMA) guidelines.

METHODS

A literature search was performed for relevant studies published from March 2020 - outbreak of the pandemic - to 30 September 2021. Seventeen articles were included in the result synthesis of this paper.

RESULTS

As part of the review, we presented the main characteristics related to AI techniques, in particular deep learning (DL), adopted in the selected articles. A survey was carried out on the type of architectures used, availability of the source code, network weights and open access datasets, use of data augmentation, use of the transfer learning strategy, type of input data and training/test datasets, and explainability.

CONCLUSION

Finally, this review highlighted the existing challenges, including the lack of large datasets of reliable COVID-19-based LUS images to test the effectiveness of DL methods and the ethical/regulatory issues associated with the adoption of automated systems in real clinical scenarios.

Key Words: Lung ultrasound; Deep learning; Neural network; COVID-19 pneumonia; Medical imaging

Core Tip: Challenging coronavirus disease 2019 (COVID-19) pandemic through the identification of effective diagnostic and prognostic tools is of outstanding importance to tackle the healthcare system burdening and improve clinical outcomes. Application of deep learning (DL) in medical lung ultrasound may offer the advantage of combining non-invasiveness and wide accessibility of ultrasound imaging techniques with higher diagnostic performance and classification accuracy. This paper overviews the current applications of DL models in medical lung ultrasound imaging in COVID-19 patients, and highlight the existing challenges associated with the effective clinical application of automated systems in the medical imaging field.

Citation: De Rosa L, L'Abbate S, Kusmic C, Faita F. Applications of artificial intelligence in lung ultrasound: Review of deep learning methods for COVID-19 fighting. *Artif Intell Med Imaging* 2022; 3(2): 42-54
URL: https://www.wjgnet.com/2644-3260/full/v3/i2/42.htm
DOI: https://dx.doi.org/10.35711/aimi.v3.i2.42

INTRODUCTION

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is a life-threatening infectious virus and its related disease (COVID-19) represents a still ongoing challenge for humans. At time of writing, over 497 million infections have been recorded worldwide including more than 6.1 million attributable deaths[1]. Despite the large number of vaccination programs introduced from the end of 2020 has represented an opportunity to minimise the risk of severe COVID-19 and death, the spread of new genetic viral variants with a higher probability of contagion has raised a renewed strong concern for either not vaccinated and vaccinated people. Thus, since the outbreak of the pandemic, research has continuously looked for a quick and reliable way to diagnose the disease, treat and monitor people affected by coronavirus.

To date, molecular test based on real time quantitative reverse transcription polymerase chain reaction (RT-qPCR) assay by nasopharyngeal swabs along with the serological antibody-detecting and antigen-detecting tests are the current accepted diagnostic tools for the conclusive diagnosis of COVID-19[2]. RT-qPCR may take up to 24 h to provide information and requires multiple tests for definitive results and, in addition, it is not relevant to assess the disease severity. Furthermore, the accuracy of molecular and serological tests remains highly dependent on timing of sample collection relative to infection, improper sampling of respiratory specimens, inadequate preservation of samples and technical errors, particularly contamination during RT-qPCR process and cross-reactivity in the immunoassay[3,4].

To complement conventional *in vitro* analytical techniques of COVID-19, biomedical imaging techniques have demonstrated great potential in clinical diagnostic evaluation by providing rapid patient assessment in the presence of high pre-test probability. Furthermore, imaging techniques are currently important in the follow-up of subjects with COVID-19[5,6]. Among the imaging techniques, chest computed tomography (CT) is considered the primary diagnostic modality and an important indicator for assessing severity and progression of COVID-19 pneumonia[7,8], although it has been reported to have limited specificity[9-11]. Indeed, the CT imaging features can overlap between COVID-19 and other viral pneumonia. Moreover, CT scanning is expensive, not easy to perform in the COVID-19 context, and multiple risks are associated with it, such as radiation exposure and cross-infection risk associated with repeated use of a CT suite[12], along with unavailability of CT in many parts of the world.

In the last few years, lung ultrasound (LUS) technique has become increasingly popular and a good option for real-time point-of-care testing, with several advantages making it a valuable tool in the fight against COVID-19[13], although it has specificity limits comparable to those of chest CT.

Ultrasound (US) is a low-cost, non-radioactive medical imaging method, particularly indicated for evaluation in pregnant women and children, which is portable to the bedside or patient's home and is easy to sterilise. Moreover, the risk of COVID-19 cross-infection can be limited by making use of

disposable ultrasound gel with a portable probe[14]. In addition, some studies indicate that LUS shows excellent performances in speed of execution and accuracy of diagnosis in case of respiratory failure [15]. Furthermore, compared with chest X-ray, LUS demonstrated higher sensitivity in detecting pneumonia[16] and similar specificity in the diagnosis of pneumothorax[15]. On the other hand, the distinctive LUS features (B-lines, consolidations, pleural thickening and rupture) observed in patients with varying severity of COVID pneumonia are similar to the features seen in patients with pneumonia of different aetiologies. Indeed, a recent review[17] on ultrasound findings of LUS in COVID-19 demonstrated that LUS has high sensitivity and reliability in ruling out lung involvement, but at the expense of low specificity. Therefore, especially in the case of low prevalence of the disease, at present LUS cannot be considered a valid gold standard in clinical practice.

Ultrasound image processing techniques have assumed great importance in recent years, with the growing experience that accurate image processing can significantly help in extracting quantitative characteristics to assess and classify the severity of diseases. Accordingly, sophisticated techniques of automated image processing, that include the use of artificial intelligence (AI) methods, have been developed and applied to assist LUS imaging in the detection of COVID-19 and make such assessment more objective and accurate. AI methods - from machine learning (ML) to deep learning (DL), indeed, aim to imitate cognitive functions and stand out in automatically recognizing complex patterns in imaging data, providing quantitative rather than qualitative assessments. The primary purpose of applying AI methods in medical imaging is to improve the visual recognition of certain features in images to produce lower-than-human error rates. Furthermore, an enhancement in LUS performance can reduce the use of more invasive and time-consuming techniques, facilitating both faster diagnosis and recognition of earlier stages of the disease[18]. To allow a quick development of highly performant AI models, a large amount of accessible and validated data to train and test AI models is a critical requirement that can be achieved, for instance, with the development of shared big data archives. Indeed, one of the most common problems associated with using limited training samples is the over-fitting of DL models. To address this issue, two main approaches can be selected: model optimization and transfer learning. These strategies significantly improve the performance of DL models. Likewise, data pre-processing and data augmentation/enhancement can be useful additional strategies[19,20].

The most common applications of DL methods in clinical imaging, and hence in medical ultrasound imaging as well, are object detection, object segmentation, and object classification[21]. The main architectures applied in current analysis are convolutional neural networks (CNNs) and recurrent neural networks (RNNs)[22]. CNNs are architectures able to work with 2D and 3D input images and RNNs recognize the image's sequential characteristics and use patterns to predict the next likely scenario[23].

Since the outbreak of the pandemic, many proposals have been made based on AI methods applied to LUS scans of COVID-19 patients. Here we propose a comprehensive systematic review of the literature on the use of AI technology, DL in particular, to aid in the fight against COVID-19.

MATERIALS AND METHODS

Study selection

A literature search to identify all relevant articles on the use of DL tools applied to LUS imaging in patients affected by COVID-19 virus was conducted.

This systematic review was carried out using the PubMed/Medline electronic database and according to the preferred reporting for systematic reviews and meta-analysis (PRISMA) guidelines[24, 25]. We performed a systematic search covering the period from March 2020 (from the outbreak of the pandemic) to 30 September 2021. The search strategy was restricted to English-language publications.

We performed an advanced research concatenating terms with Boolean operators. In particular, search words and key terms used in the search included ("lung ultrasound" OR "lus") AND ("COVID-19" OR "coronavirus" OR "SARS-CoV2") AND ("artificial intelligence" OR "deep learning" OR "neural networks" OR "CNN").

Eligibility criteria

The inclusion criteria were: Studies that include COVID-19 patients with LUS acquisitions and developed or tested DL-based algorithms on LUS images or on features extracted from the images; No restriction on the ground truth adopted to analyse the presence/absence of COVID-19 and/or the severity of lung disease (*e.g.*, PCR, visual evaluation of video/images and score assignment by expert clinicians); No restriction on the type of DL architecture used in the studies. Studies on paediatric population were excluded. Studies were restricted to peer reviewed articles and conference proceedings. However, the following publication types were excluded: reviews and conference abstracts.

Data extraction and analysis

Two investigators (DRL and FF) screened the articles independently. Disagreement between reviewers was resolved by consensus *via* discussion. The reasons for the exclusion of some trials are described in the Results section. Publications by the same research group or by different groups using the same dataset were included in the analysis. After the selection of the articles, we collected the following characteristics: First author's surname, date of publication, sample size, general characteristics of the study populations, AI techniques used, validation methods and main results obtained. The study selection process is presented in Figure 1.

RESULTS

Search results

Twenty-four articles resulted after querying the database and screened for eligibility (Figure 1). Of the 24 articles, we discarded four references as review papers. After examining the titles and abstracts, we excluded five articles: one manuscript did not include DL methods applied on US imaging, three papers were not based on AI and DL approaches, and one article was focused on the paediatric population. Moreover, two additional papers, retrieved from the checking of references of the eligible articles, were included. Finally, 17 articles[26-42] were selected for full-text screening and included in our analysis (Table 1 and 2). The following part of the section provides a concise overview of the studies' main features.

Dataset and source code availability

Authors of seven[27-30,33,39,40] of the seventeen selected articles (41.2%) extrapolated their datasets from the free access LUS database acquired by point-of-care ultrasound imaging and made available firstly by Born *et al*[30]. Instead, an Italian group firstly introduced the Italian COVID-19 Lung Ultrasound DataBase (ICLUS-DB)[38], which is accessible upon mandatory request to the authors, and that was used in two other studies[32,37]. Noteworthy, Roy *et al*[38] have created a platform through which physicians can access algorithms, upload their data and see the algorithm's evaluation of the data.

Besides dataset open access, access to the code for the neural network is also important to reproduce results and compare performances. Seven articles[26-30,32,38] (41.2%) made the source code implementing the proposed DL architecture available for download from the Git-hub repository.

Single-frame/multi-frames or video based architecture

In the majority of the selected papers, DL architectures work with single frame images as input and only three publications[29,34,41] (17.6%) report DL architectures based on image sequences (*i.e.*, video). However, six studies[28,30,32,37-39] (35.3%), despite adopting a DL architecture designed to perform single-frame classification, also propose additional methods to fulfil video-based classification. In particular, Roy *et al*[38] proposed an aggregation layer system of frame-level scores to produce predictions on LUS videos and Mento *et al*[37] proposed an alternative video-based classification using a threshold-based system on the frame-level scores obtained from DL architecture.

Other authors[32] adopted a Long Short-Term Memory (LSTM) system, which has been used to exploit temporal relationships between multiple frames by taking long time series as input, over performing their results obtained by CNN without LSTM.

Finally, Xue *et al*[42] applied AI models for patient-level assessment of severity using a final module across the entire architecture that works with ML rather than DL systems.

Test strategy of DL models

The proposed DL models have been tested on a database entirely independent from the training database in seven articles[26,35-39,42] (41.2%); five-fold and ten-fold cross-validation techniques were applied in nine[27-34,40] (52.9%) and one[41] (5.9%) studies, respectively. Among the papers that tested DL models on an independent database, the percentage of data used for the testing ranged from 33%[35] to 20%[38] and 10%[26,36] of the overall data. Born *et al*[29], alongside the five-fold cross-validation technique in the training/test phase of the DL model, also used an independent validation dataset made-up of 31 videos (28 convex and 3 linear probes) from six patients. Indeed, Roy *et al*[38], for instance, used 80 videos/10709 frames out of the total 277 videos/58924 frames to test their DL model.

In all studies, the splitting of data between training set and test set was performed either at the patient-level or at the video-level. Thus, all the frames of a single video clip belonged either to the training or to the test set.

Data augmentation

Twelve (70.6%) research groups extended their LUS database by augmentation. The main strategies for data augmentation applied to LUS images were: Horizontal/vertical flipping[26,27,29,30,32,33,36,38-40,

Table 1 General characteristics of the studies included in the analysis (part I)

Ref.	Publication date	Journal	Sample size[1], N° pts/videos/images	Subjects	Main results
Arntfield *et al* [26]	22/02/2021	BMJ Open	243/612/121k	COVID +, COVID -, HPE	Overall Acc = 0.978AUC = 1/0.934/1 for COVID +, COVID -, HPE
Awatshi *et al* [27]	23/03/2021	IEEE Trans Ultrason Ferroelectr Freq Control	-/64/1.1k	COVID +, Healthy, PN	5-fold validation: Acc = 0.829
Barros *et al*[28]	14/08/2021	Sensors	131/185/-	COVID +, PN bacterial, Healthy	Best model (Xception+LSTM): Acc = 0.93 – Se = 0.97
Born *et al*[29]	12/01/2021	Applied Sciences	216/202/3.2k	COVID +, Healthy, PN	External validation: Se = 0.806 – Sp = 0.962
Born *et al*[30]	24/01/2021	ISMB TransMed	-/64/1.1k	COVID +, Healthy, PN	Overall Acc = 0.89Binarization COVID y/n: Se = 0.96 – Sp = 0.79 – F1score = 0.92
Chen *et al*[31]	29/06/2021	IEEE Trans Ultrason Ferroelectr Freq Control	31/45/1.6k	COVID-19 PN	5-fold validation: Acc = 0.87
Dastider *et al* [32]	20/02/2021	Comput Biol Med	29/60/14.3k	COVID-19 PN	Independent data validation: Acc = 0.677 – Se = 0.677 – Sp = 0.768 – F1score = 0.666
Diaz Escobar *et al*[33]	13/08/2021	PLos One	216/185/3.3k	COVID +, PN bacterial, Healthy	Best model (InceptionV3): Acc = 0.891 – AUC = 0.971
Erfanian Ebadi *et al*[34]	04/08/2021	Inform Med Unlocked	300/1.5k/288k	COVID +, PN	5-fold validation: Acc = 0.90 – PP=0.95
Hu *et al*[35]	20/03/2021	BioMed Eng OnLine	108/-/5.7k	COVID +	COVID detection: Acc = 0.944 – PP = 0.823 – Se = 0.763 – Sp=0.964
La Salvia *et al* [36]	03/08/2021	Comput Biol Med	450/5.4k/> 60k	Hospitalised COVID-19	External validation (ResNet50): Acc = 0.979 – PP=0.978 – F1score = 0.977 – AUC = 0.998
Mento *et al*[37]	27/05/2021	J Acoust Soc Am	82/1.5k/315k	COVID-19 confirmed	% Agreement DL and LUS = 96%
Roy *et al*[38]	14/05/2020	IEEE Trans	35/277/58.9k	COVID-19 confirmed, COVID-19 suspected, Healthy	Segmentation: Acc = 0.96 – DICE = 0.75
Sadik *et al*[39]	09/07/2021	Health Inf Sci Syst	-/123/41.5k	COVID +, PN, Healthy	COVID y/n (VGG19+SpecMen): PP = 0.81 – F1score = 0.89
Muhammad *et al*[40]	25/02/2021	Information Fusion	121 videos + 40 frames	COVID +, PN bacterial, Healthy	Overall: Acc = 0.918 – PP = 0.925
Tsai *et al*[41]	08/03/2021	Phys Med	70/623/99.2k	Healthy, Pleural effusion pts	Pleural effusion detection:Acc = 0.924
Xue *et al*[42]	20/01/2021	Med Image Anal	313/-/6.9k	COVID-19 confirmed	4-level and binary disease severity:Acc = 0.75 and Acc = 0.85

[1]k: Indicates × 10^3.

pts: Patients; HPE: Hydrostatic pulmonary edema; PN: Pneumonia; Acc: Accuracy; Se: Sensitivity; Sp: Specificity; AUC: Area under the curve; PP: Precision; DL: Deep learning; LUS: Lung ultrasound.

42], bidirectional arbitrary rotation[26,27,29,30,32,33,35,38-40,42], horizontal and vertical shift[30,32,38,39,42]; filtering, colour transformation, adding salt and pepper noise, Gaussian noise[36,38,42], normalisation of grey levels' intensity[38]. Although proposed by all the authors, only seven papers[26,29,30,32,33,38,40] provided details on the amplitude of image rotation. In particular, Dastider *et al*[32] applied rotations in the range of 0 ± 360 degrees, while other authors have limited image rotations to 10 degrees [26,29,30,33], ± 15 degrees[38] and ± 20 degrees[40], respectively. The remaining five papers[28,31,34,37,41] (29.4%) did not perform data augmentation.

Explainability

Among the selected articles, tools for interpreting the network output were provided in twelve studies (70.6%), whereas in the remaining five (29.4%) the DL algorithms' outcomes were proposed as black box systems. The majority of papers[26-29,32,35,36,38,40] reported the Gradient-weighted Class Activation

Table 2 General characteristics of the studies included in the analysis (part II)

Ref.	DL architecture	Input of DL models	Available dataset	Available code	Pre-trained/TL	Test independent	Data Augmentation	Explainability
Arntfield *et al* [26]	CNN	SF	No	Yes (on github)	Yes	Yes	Yes	Yes
Awatshi *et al* [27]	CNN	SF	No	Yes (on github)	Yes	No (five-fold)	Yes	Yes
Barros *et al*[28]	CNN+LSTM	SF	Yes	Yes (on github)	Yes	No(five-fold)	No	Yes
Born *et al*[29]	3D CNN	MF	Yes	Yes (on github)	Yes	No(five-fold)	Yes	Yes
Born *et al*[30]	CNN	SF	Yes	Yes (on github)	Yes	No(five-fold)	Yes	No
Chen *et al*[31]	MLFCNN	SF	No	Yes (on github)	No	No(five-fold)	No	No
Dastider *et al* [32]	CNN+LSTM	SF	No	Yes (on github)	Yes	No(five-fold)	Yes	Yes
Diaz Escobar *et al*[33]	CNN	SF	No	No	Yes	No(five-fold)	Yes	No
Erfanian Ebadi *et al*[34]	3D CNN	MF	No	Yes (on github)	Yes	No(five-fold)	No	Yes
Hu *et al*[35]	CNN + MCRF	SF	No	No	Yes	Yes	Yes	Yes
La Salvia *et al* [36]	CNN	SF	No	No	Yes	Yes	Yes	Yes
Mento *et al*[37]	CNN+ STN	SF	No	No	No	-	No	No
Roy *et al*[38]	CNN+ STN	SF	Yes (on request)	Yes (on github)	No	Yes	Yes	Yes
Sadik *et al*[39]	CNN	SF	No	No	Yes	Yes	Yes	Yes
Muhammad *et al*[40]	CNN	SF	Yes	No	No	No(five-fold)	Yes	Yes
Tsai *et al*[41]	CNN+ STN	MF	No	No	Yes	No(ten-fold)	No	No
Xue *et al*[42]	CNN	SF	No	No	No	Yes	Yes	Yes

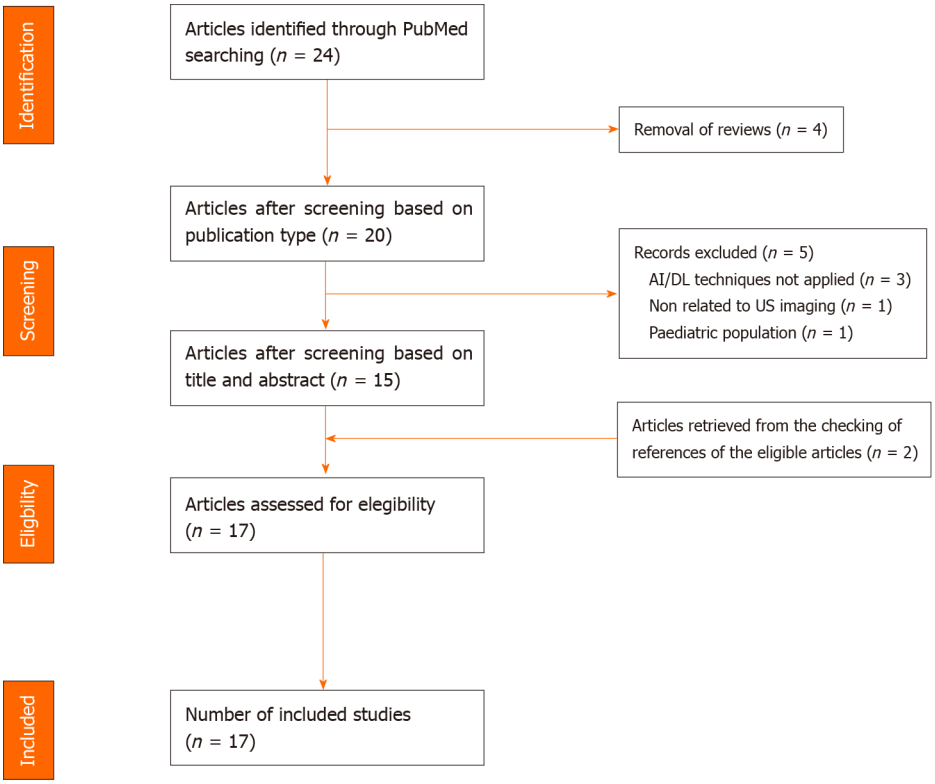
CNN: Convolutional neural network; LSTM: Long short-term memory; MCRF: Multimodal channel and receptive field; MLFCNN: Multi-layer fully connected neural network; STN: Spatial transformer network; SF: Single-frame; MF: Multi-frame; DL: Deep learning; TL: Transfer learning.

Mapping (Grad-CAM) as the preferred explainability tool. Grad-CAM uses gradients to create a location map to highlight the region of interest of the images[43]. Instead, Sadik *et al*[39] used a colormap jet to visualise a heat map overlay to US images; Erfanian Ebadi *et al*[34] adopted an activation map system to detect and segment features in LUS scans. Furthermore, one study[42] showed LUS images with overlaid colormaps to indicate the segmentation zone of ultrasound according to the different severity. Roy *et al*[38], differently, provided an ultrasound colormap overlay on the LUS frame/video and used four colours to distinguish the different classes of disease severity recognized by DL architecture.

Clinical use

Most of the selected papers applied the AI system to diagnose COVID-19 and/or discriminate between COVID-19 and other lung diseases (such as bacterial pneumonia)[26-30,33,34,39,40]. The first approach using DL architecture for automatic differential diagnosis of COVID-19 from LUS data was POCOVID-Net[30].

However, a fair number of studies have focused on assessing the severity of COVID-19[31,32,35-38,42]. In particular, a disease severity score is assigned to the single image according to some characteristics visible in the image pattern. Most of the articles used four severity classes by assigning a score to the single frame from 0 to 3[31,32,35-38], as defined by Soldati *et al*[44]. Xue *et al*[42] proposed a classification in five classes of pneumonia severity (score from 0 to 4) along with a binary severe/non-severe classification. Furthermore, these authors used the DL technology exclusively to implement a segmentation phase based on a VGG network, while the classification phase still employed a more traditional, features-based machine learning approach. Finally, La Salvia *et al*[36] proposed a classi-



DOI: 10.12998/wjcc.v3.i2.42

Figure 1 Flow diagram of systematic identification, screening, eligibility and inclusion of publications that applied deep learning methods to lung ultrasound imaging in coronavirus disease 2019 patients. AI: Artificial intelligence; DL: Deep learning; US: Ultrasound.

fication based on three severity classes and a modified version considering a seven-classes scenario.

Furthermore, Arntfield *et al*[26] showed that their network was able to recognize pathological pattern in LUS images with higher sensitivity than sonographers; whilst an InceptionV3 network proposed by Diaz-Escobar *et al*[33] was able to discriminate COVID-19 pneumonia from healthy lung and other bacterial pneumonia with an accuracy of 89.1% and an area under the ROC curve of 97.1%.

Curiously, one of the eligible papers[41] did not include confirmed cases of COVID-19 patients. The authors' aim was to design an algorithm capable of identifying the presence of pleural effusion. However, we have included this work in our systematic review, because small pleural effusions are rarely reported in COVID-19 patients. Therefore, the detection of pneumonia with pleural effusion can help rule out the hypothesis of COVID-19 disease.

Transfer learning and DL architecture

From our analysis, it emerged that most of the studies have proposed convolutional neural networks (CNNs) as DL models to generate screening systems for COVID-19. In particular, all publications with the exception of one[31] used the CNN network. Conversely, Chen *et al*[31] developed a multi-layer fully connected neural network for scoring LUS images in assessing the severity of COVID-19 pneumonia.

Among the DL systems included in this review, most of them were generated starting from DL architectures already proposed for other tasks[26-30,32-36,39,42], suitably modified and trained for new tasks. Furthermore, many works compared the results of their architectures with those obtained using existing and well-known architectures[27-30,32,33,35,38-40]. In particular, the following DL architectures were adapted to fulfil the requirements of LUS analysis to assist in COVID-19 detection and/or assessment of the severity of the lung disease, or just to compare their performances: VGG-19 [28,33,39] and VGG-50[28-30,33]; Xception[26,28,39]; ResNet 50[27,33,36,40]; NasNetMobile[27,29,39]; DenseNet[32,39].

More in detail, Awasthi *et al*[27] proposed Mini-COVIDNet, a modified MobileNet model belonging to the CNN's networks family and originally developed for detecting objects in mobile applications[45]. Barros *et al*[28], along with their proposed DL model, also investigated the impact of using different pre-trained CNN architectures in extracting spatial features that were successively classified by a LSTM model. Finally, Born *et al*[29] derived their DL video-based models from a model that was pre-trained on lung CT scans[46].

All aforementioned architectures are pre-trained on ImageNet[47].

Sample size

Partly due to the recent outbreak of the pandemic and to the difficulty of having standardised high quality archives of US images, only few of the selected studies relied on a large dataset in terms of enrolled patients. Six papers (35.3%) reported a sample size greater than 200 subjects (namely, 243, 216, 216, 300, 450 and 313 in references[26,29,33,34,36,42] respectively).

However, despite the relatively low number of subjects, the total number of LUS videos reaches up to 5400 in one study[36], with an average equal to 1589 videos[26,29,33,34,36]. Among the studies carried out on a low sample size, Dastider *et al*[32] included 29 patients and 60 videos, whilst 35 patients/45 videos and 35 patients/277 videos were analysed in references Chen *et al*[31] and Roy *et al*[38], respectively. However, it should be noted that Roy *et al*[38] published their work at the beginning of the COVID-19 pandemic, when the total number of COVID-19 patients was still relatively limited. In the paper by Xue *et al*[42], the number of frames/video was not reported.

DISCUSSION

The paper reviews the different DL techniques able to work with LUS images in assisting the diagnosis and/or prognosis of the COVID-19 disease published since the outbreak of the pandemic. In the selected documents, the use of DL systems aimed to achieve an accuracy comparable to or better than clinical standards to provide a faster diagnosis and/or follow-up in COVID-19 patients.

Most of the papers present pre-trained DL architectures[26-30,32-36,39,42] that were modified and adapted to new data. This approach is also known as transfer learning (TL) technique - *i.e.*, a training strategy for new DL models with reduced datasets. The network is pre-trained on a very large dataset, such as ImageNet, with millions of images intentionally created to facilitate the training of DL models, focusing on image classification and object location/detection tasks[48]. Indeed, deeper models are difficult to train and provide inconsistent performances when trained on a limited amount of data[49]. Therefore, most of the studies based on DL systems to classify COVID-19 images appropriately use the TL strategy as large datasets of US images from COVID-19 patients are not yet easily available, partly because the coronavirus disease is a relatively recent concern.

Furthermore, most of the proposed systems shared the same design, *i.e.*, CNN's architectures. CNNs have several applications in medical imaging – among others, image segmentation and object detection [50]. However, CNNs are particularly suited for image classification problems[51] and, consequently, represent an optimal solution for the classification of the disease severity from US images.

To date, one of the main challenges faced by DL architectures applied to LUS images of COVID-19 patients are the limited datasets in the available databases. This problem could benefit from creating open access databases that collect large amounts of data from multiple centres. In some of the selected studies, a first attempt to overcome this issue is evident, with particular emphasis on the work by Born *et al*[30], the authors who first collected a free access dataset of lung images from healthy controls and patients affected by COVID-19 or other pneumonia.

The development of public and multicentre platforms would guarantee the collection of a continuously growing amount of data, large and highly heterogeneous, suited for the training and testing of new DL applications in medical imaging, both in the COVID-19 and LUS field. Furthermore, this would allow an easier comparison of performances among DL models proposed in different studies. However, alternative approaches are often used in the testing phase that do not require the use of independent data sets to evaluate the performance of the model in the event of a limited number of images available. Among these, the k-fold cross-validation is a statistical method used to evaluate the ability of ML models to generalise to previously unseen data. Despite being widely used in ML models, the k-fold cross validation approach is less reliable than tests performed using an external dataset; the latter is always preferable to test model's ability to adapt properly to new, previously unseen data.

Data augmentation techniques are an alternative strategy to overcome the issue of the limited amounts of data, largely adopted in practice. These techniques generate different versions of a real dataset artificially to both increase its size and the power of model's generalisation. Despite the great advantage in increasing data to feed DL architectures, data augmentation techniques should be used with awareness, as some geometric transformations could be unrealistic when applied to LUS images (*e.g.*, angles of rotations greater than 30°). In the field of DL applied to medical imaging, the use of architectures designed to work with 3D images is another interesting challenge. Indeed, a DL system that operates with 3D data input usually requires a larger amount of data for training, as a 3D network contains a parameters' number that is orders of magnitude greater than a 2D network. This could significantly increase the risk of overfitting, especially in the case of limited dataset availability. In addition, the training on large amounts of data requires high computational costs associated with memory and performance requirements of the tools used. LUS images are usually recorded in the form of videoclips (2D + time) and can be assimilated to 3D data. Exploitation of dynamic information naturally embedded in image sequences has proven very important in the analysis of lung echoes. In

particular, changes induced by COVID-19 viral pneumonia are better detectable in LUS through the analysis of multi-frames acquisition due to its ability in capturing dynamic features, *e.g.*, pleural sliding movements and generation of B-line artefacts[44].

Regardless of the data format (*i.e.*, 3D, 2D or 2D+time images), the labelling of ground truth data is required in supervised DL applications and should be provided by skilled medical professionals. However, it is a time-consuming activity, in particular in the 2D approach that is characterised by a high number of samples.

Indeed, some authors demonstrated that the performance in pleural effusion classification on LUS images obtained with the video-based approach was comparable to that obtained with frame-based analysis, despite a significant reduction in labelling effort[41]. Furthermore, Kinetics-I3D network was able to classify LUS video sequences with great accuracy and efficiency[34]. On the other hand, the video-based approach has also revealed a reduced accuracy in patients classification with respect to the single frame analysis; however, this could be explained by the relatively reduced number of available LUS clips[29].

Extending the use of DL architectures beyond multi-frame analysis with respect to single 2D images is highly desirable. In particular, these methods could be effectively used to assign a patient-level disease severity score. In fact, this information plays a key role in the selection of treatment, monitoring of disease progression and management of medical resources (*e.g.*, mechanical ventilator needed).

Code availability is another very critical issue in applications of AI in medical imaging. Indeed, the lack of ability to reproduce the training of the proposed DL models or to test these models on new US images is a rather widespread problem. Often, authors do not provide access to either the source code used to train NNs or the final weight of the trained network. On the other hand, the availability of this information would greatly facilitate the diffusion of new AI systems in the clinical setting.

DL systems are often presented as black boxes - *i.e.*, they produce a result without providing a clear understanding in "human terms" of how it was obtained. The black-box nature of the algorithms has restricted their clinical use until now. Consistently, the explainability - *i.e.*, making clear and understandable the features that influence the decisions of a DL model - is a critical point to guarantee a safe, ethical, and reliable use of AI. Especially in medical imaging applications, explainability is very important as it gives the opportunity to highlight regions of the image containing the visual features that are critical for the diagnosis. Gradient-weighted Class Activation Mapping (Grad-CAM) is a promising technique for producing "visual explanations" of decisions taken from a large class of CNN-based models, making their internal behaviour more understandable, thus partially overcoming the black-box problem. The basic idea is to produce a rough localization map that highlights the key regions in the image that have a major effect on customization of network parameters, thus maximally contributing to the prediction of outcomes[43].

These maps visualised areas using a blue-to-red scale, with the highest/lowest contribution to the class prediction operated by the model. The clinical use of DL systems is a crucial issue. One of the major current limitations of LUS imaging in COVID patients is the specificity. Focusing the design of DL systems to overcome this limit could really represent a benefit in the clinical setting.

Along this line, some of the included studies tested the agreement between physicians' ability to classify COVID-19 patients and that proposed by neural networks. Furthermore, this finding suggests that the automated system can capture some features (biomarkers) in US images that are not clearly visible to the human eye.

Finally, another important issue to mention is the use of the quantitative evaluation indicators and the analysis of the benchmarking techniques adopted to evaluate the effectiveness of the proposed methods. Unfortunately, the tools examined in the selected manuscripts had very heterogeneous targets (Table 1, Main results column), ranging from diagnostic to prognostic purposes or assessment of disease severity. This dispersion of intent and the few articles published in the literature at present make any comparison or analysis very difficult.

CONCLUSION

The studies analysed in this article have shown that DL systems applied to LUS images for the diagnosis/prognosis of COVID-19 disease have the potential to provide significant support to the medical community. However, there are a number of challenges to overcome before AI systems can be regularly employed in the clinical setting. On the one hand, the critical issues related to the availability of high-quality databases with large sample size of lung images/videos of COVID-19 patients and free access to datasets must be addressed. On the other hand, existing concerns about the methodological transparency (*e.g.*, explainability and reproducibility) of DL systems and the regulatory/ethical and cultural issues that the clinical use of AI methods raise must be resolved. Finally, a closer collaboration between the communities of informatics/engineers and medical professionals is desirable to facilitate the outcome of adequate guidelines for the use of DL in US pulmonary imaging and, more generally, in medical imaging.

ARTICLE HIGHLIGHTS

Research background

The current coronavirus disease 2019 (COVID-19) pandemic crisis has highlighted the need for biomedical imaging techniques in rapid clinical diagnostic evaluation of patients. Furthermore, imaging techniques are currently important in the follow-up of subjects with COVID-19. The lung ultrasound technique has become increasingly popular and is considered a good option for real-time point-of-care testing, although it has specificity limits comparable to those of chest computed tomography.

Research motivation

The application of artificial intelligence, and of deep learning in particular, in medical pulmonary ultrasound can offer an improvement in diagnostic performance and classification accuracy to a non-invasive and low-cost technique, thus implementing its diagnostic and prognostic importance to COVID-10 pandemic.

Research objectives

This review presents the state of the art of the use of artificial intelligence and deep learning techniques applied to lung ultrasound in COVID-19 patients.

Research methods

We performed a literature search, according to preferred reporting items of systematic reviews and meta-analysis guidelines, for relevant studies published from March 2020 - to 30 September 2021 on the use of deep learning tools applied to lung ultrasound imaging in COVID-19 patients. Only English-language publications were selected.

Research results

We surveyed the type of architectures used, availability of the source code, network weights and open access datasets, use of data augmentation, use of the transfer learning strategy, type of input data and training/test datasets, and explainability.

Research conclusions

Application of deep learning systems to lung ultrasound images for the diagnosis/prognosis of COVID-19 disease has the potential to provide significant support to the medical community. However, there are critical issues related to the availability of high-quality databases with large sample size and free access to datasets.

Research perspectives

Close collaboration between the communities of computer scientists/engineers and medical professionals could facilitate the outcome of adequate guidelines for the use of deep learning in ultrasound lung imaging.

FOOTNOTES

Author contributions: Kusmic C and Faita F designed the research study; Faita F and De Rosa L collected and analysed the references mentioned in the review; De Rosa L wrote the initial draft; Kusmic C, Faita F and L'Abbate S revised and edited the manuscript; all authors have read and approve the final manuscript.

Conflict-of-interest statement: The authors declare that they have no conflict of interest.

PRISMA 2009 Checklist statement: The authors have read the PRISMA 2009 Checklist, and the manuscript was prepared and revised according to the PRISMA 2009 Checklist.

Country/Territory of origin: Italy

ORCID number: Laura De Rosa 0000-0002-6030-8483; Serena L'Abbate 0000-0002-4462-0645; Claudia Kusmic 0000-0002-8064-8311; Francesco Faita 0000-0002-6201-1843.

S-Editor: Liu JH
L-Editor: A
P-Editor: Liu JH

REFERENCES

1 **World Health Organization**. (2022, April 13). Coronavirus (COVID-19) Dashboard. Available from: https://covid19.who.int
2 **Li Z**, Yi Y, Luo X, Xiong N, Liu Y, Li S, Sun R, Wang Y, Hu B, Chen W, Zhang Y, Wang J, Huang B, Lin Y, Yang J, Cai W, Wang X, Cheng J, Chen Z, Sun K, Pan W, Zhan Z, Chen L, Ye F. Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. *J Med Virol* 2020; **92**: 1518-1524 [PMID: 32104917 DOI: 10.1002/jmv.25727]
3 **Yang Y**, Yang M, Yuan J, Wang F, Wang Z, Li J, Zhang M, Xing L, Wei J, Peng L, Wong G, Zheng H, Wu W, Shen C, Liao M, Feng K, Yang Q, Zhao J, Liu L, Liu Y. Laboratory Diagnosis and Monitoring the Viral Shedding of SARS-CoV-2 Infection. *Innovation (N Y)* 2020; **1**: 100061 [PMID: 33169119 DOI: 10.1016/j.xinn.2020.100061]
4 **Mardian Y**, Kosasih H, Karyana M, Neal A, Lau CY. Review of Current COVID-19 Diagnostics and Opportunities for Further Development. *Front Med (Lausanne)* 2021; **8**: 615099 [PMID: 34026773 DOI: 10.3389/fmed.2021.615099]
5 **Hoffmann T**, Bulla P, Jödicke L, Klein C, Bott SM, Keller R, Malek N, Fröhlich E, Göpel S, Blumenstock G, Fusco S. Can follow up lung ultrasound in Coronavirus Disease-19 patients indicate clinical outcome? *PLoS One* 2021; **16**: e0256359 [PMID: 34432835 DOI: 10.1371/journal.pone.0256359]
6 **Martini K**, Larici AR, Revel MP, Ghaye B, Sverzellati N, Parkar AP, Snoeckx A, Screaton N, Biederer J, Prosch H, Silva M, Brady A, Gleeson F, Frauenfelder T; European Society of Thoracic Imaging (ESTI), the European Society of Radiology (ESR). COVID-19 pneumonia imaging follow-up: when and how? *Eur Radiol* 2021 [PMID: 34713328 DOI: 10.1007/s00330-021-08317-7]
7 **Bernheim A**, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, Diao K, Lin B, Zhu X, Li K, Li S, Shan H, Jacobi A, Chung M. Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection. *Radiology* 2020; **295**: 200463 [PMID: 32077789 DOI: 10.1148/radiol.2020200463]
8 **Lyu P**, Liu X, Zhang R, Shi L, Gao J. The Performance of Chest CT in Evaluating the Clinical Severity of COVID-19 Pneumonia: Identifying Critical Cases Based on CT Characteristics. *Invest Radiol* 2020; **55**: 412-421 [PMID: 32304402 DOI: 10.1097/RLI.0000000000000689]
9 **Raptis CA**, Hammer MM, Short RG, Shah A, Bhalla S, Bierhals AJ, Filev PD, Hope MD, Jeudy J, Kligerman SJ, Henry TS. Chest CT and Coronavirus Disease (COVID-19): A Critical Review of the Literature to Date. *AJR Am J Roentgenol* 2020; **215**: 839-842 [PMID: 32298149 DOI: 10.2214/AJR.20.23202]
10 **Inui S**, Gonoi W, Kurokawa R, Nakai Y, Watanabe Y, Sakurai K, Ishida M, Fujikawa A, Abe O. The role of chest imaging in the diagnosis, management, and monitoring of coronavirus disease 2019 (COVID-19). *Insights Imaging* 2021; **12**: 155 [PMID: 34727257 DOI: 10.1186/s13244-021-01096-1]
11 **Xu B**, Xing Y, Peng J, Zheng Z, Tang W, Sun Y, Xu C, Peng F. Chest CT for detecting COVID-19: a systematic review and meta-analysis of diagnostic accuracy. *Eur Radiol* 2020; **30**: 5720-5727 [PMID: 32415585 DOI: 10.1007/s00330-020-06934-2]
12 **Self WH**, Courtney DM, McNaughton CD, Wunderink RG, Kline JA. High discordance of chest x-ray and computed tomography for detection of pulmonary opacities in ED patients: implications for diagnosing pneumonia. *Am J Emerg Med* 2013; **31**: 401-405 [PMID: 23083885 DOI: 10.1016/j.ajem.2012.08.041]
13 **Smith MJ**, Hayward SA, Innes SM, Miller ASC. Point-of-care lung ultrasound in patients with COVID-19 - a narrative review. *Anaesthesia* 2020; **75**: 1096-1104 [PMID: 32275766 DOI: 10.1111/anae.15082]
14 **Akl EA**, Blažić I, Yaacoub S, Frija G, Chou R, Appiah JA, Fatehi M, Flor N, Hitti E, Jafri H, Jin ZY, Kauczor HU, Kawooya M, Kazerooni EA, Ko JP, Mahfouz R, Muglia V, Nyabanda R, Sanchez M, Shete PB, Ulla M, Zheng C, van Deventer E, Perez MDR. Use of Chest Imaging in the Diagnosis and Management of COVID-19: A WHO Rapid Advice Guide. *Radiology* 2021; **298**: E63-E69 [PMID: 32729811 DOI: 10.1148/radiol.2020203173]
15 **Walden A**, Smallwood N, Dachsel M, Miller A, Stephens J, Griksaitis M. Thoracic ultrasound: it's not all about the pleura. *BMJ Open Respir Res* 2018; **5**: e000354 [PMID: 30305907 DOI: 10.1136/bmjresp-2018-000354]
16 **Amatya Y**, Rupp J, Russell FM, Saunders J, Bales B, House DR. Diagnostic use of lung ultrasound compared to chest radiograph for suspected pneumonia in a resource-limited setting. *Int J Emerg Med* 2018; **11**: 8 [PMID: 29527652 DOI: 10.1186/s12245-018-0170-2]
17 **Gil-Rodríguez J**, Pérez de Rojas J, Aranda-Laserna P, Benavente-Fernández A, Martos-Ruiz M, Peregrina-Rivas JA, Guirao-Arrabal E. Ultrasound findings of lung ultrasonography in COVID-19: A systematic review. *Eur J Radiol* 2022; **148**: 110156 [PMID: 35078136 DOI: 10.1016/j.ejrad.2022.110156]
18 **Langlotz CP**, Allen B, Erickson BJ, Kalpathy-Cramer J, Bigelow K, Cook TS, Flanders AE, Lungren MP, Mendelson DS, Rudie JD, Wang G, Kandarpa K. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology* 2019; **291**: 781-791 [PMID: 30990384 DOI: 10.1148/radiol.2019190613]
19 **Abu Anas EM**, Seitel A, Rasoulian A, St John P, Pichora D, Darras K, Wilson D, Lessoway VA, Hacihaliloglu I, Mousavi P, Rohling R, Abolmaesumi P. Bone enhancement in ultrasound using local spectrum variations for guiding percutaneous scaphoid fracture fixation procedures. *Int J Comput Assist Radiol Surg* 2015; **10**: 959-969 [PMID: 25847667 DOI: 10.1007/s11548-015-1181-6]
20 **Krizhevsky A**, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2017; **60**: 84-90 [DOI: 10.1145/3065386]

21 **Currie GM**. Intelligent Imaging: Artificial Intelligence Augmented Nuclear Medicine. *J Nucl Med Technol* 2019; **47**: 217-222 [PMID: 31401616 DOI: 10.2967/jnmt.119.232462]

22 **Liu S**, Wang Y, Yang X, Lei B, Liu L, Li SX, Ni D, Wang T. Deep learning in medical ultrasound analysis: A review. Engineering 2019; 5: 261-275 [DOI: 10.1016/j.eng.2018.11.020]

23 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

24 **Moher D**, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535 [PMID: 19622551 DOI: 10.1136/bmj.b2535]

25 **Liberati A**, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339**: b2700 [PMID: 19622552 DOI: 10.1136/bmj.b2700]

26 **Arntfield R**, VanBerlo B, Alaifan T, Phelps N, White M, Chaudhary R, Ho J, Wu D. Development of a convolutional neural network to differentiate among the etiology of similar appearing pathological B lines on lung ultrasound: a deep learning study. *BMJ Open* 2021; **11**: e045120 [PMID: 33674378 DOI: 10.1136/bmjopen-2020-045120]

27 **Awasthi N**, Dayal A, Cenkeramaddi LR, Yalavarthy PK. Mini-COVIDNet: Efficient Lightweight Deep Neural Network for Ultrasound Based Point-of-Care Detection of COVID-19. *IEEE Trans Ultrason Ferroelectr Freq Control* 2021; **68**: 2023-2037 [PMID: 33755565 DOI: 10.1109/TUFFC.2021.3068190]

28 **Barros B**, Lacerda P, Albuquerque C, Conci A. Pulmonary COVID-19: Learning Spatiotemporal Features Combining CNN and LSTM Networks for Lung Ultrasound Video Classification. *Sensors (Basel)* 2021; **21** [PMID: 34450928 DOI: 10.3390/s21165486]

29 **Born J**, Wiedemann N, Cossio M, Buhre C, Brändle G, Leidermann K, Aujayeb A, Moor M, Rieck B, et al Accelerating detection of lung pathologies with explainable ultrasound image analysis. Appl Sci 2021; 11: 672 [DOI: 10.3390/app11020672]

30 **Born J**, Brändle G, Cossio M, Disdier M, Goulet J, Roulin J, Wiedemann N. POCOVID-Net: Automatic detection of COVID-19 from a new lung ultrasound dataset (POCUS). ISMB TransMed COSI 2020 2021. Preprint

31 **Chen J**, He C, Yin J, Li J, Duan X, Cao Y, Sun L, Hu M, Li W, Li Q. Quantitative Analysis and Automated Lung Ultrasound Scoring for Evaluating COVID-19 Pneumonia With Neural Networks. *IEEE Trans Ultrason Ferroelectr Freq Control* 2021; **68**: 2507-2515 [PMID: 33798078 DOI: 10.1109/TUFFC.2021.3070696]

32 **Dastider AG**, Sadik F, Fattah SA. An integrated autoencoder-based hybrid CNN-LSTM model for COVID-19 severity prediction from lung ultrasound. *Comput Biol Med* 2021; **132**: 104296 [PMID: 33684688 DOI: 10.1016/j.compbiomed.2021.104296]

33 **Diaz-Escobar J**, Ordóñez-Guillén NE, Villarreal-Reyes S, Galaviz-Mosqueda A, Kober V, Rivera-Rodriguez R, Lozano Rizk JE. Deep-learning based detection of COVID-19 using lung ultrasound imagery. *PLoS One* 2021; **16**: e0255886 [PMID: 34388187 DOI: 10.1371/journal.pone.0255886]

34 **Erfanian Ebadi S**, Krishnaswamy D, Bolouri SES, Zonoobi D, Greiner R, Meuser-Herr N, Jaremko JL, Kapur J, Noga M, Punithakumar K. Automated detection of pneumonia in lung ultrasound using deep video classification for COVID-19. *Inform Med Unlocked* 2021; **25**: 100687 [PMID: 34368420 DOI: 10.1016/j.imu.2021.100687]

35 **Hu Z**, Liu Z, Dong Y, Liu J, Huang B, Liu A, Huang J, Pu X, Shi X, Yu J, Xiao Y, Zhang H, Zhou J. Evaluation of lung involvement in COVID-19 pneumonia based on ultrasound images. *Biomed Eng Online* 2021; **20**: 27 [PMID: 33743707 DOI: 10.1186/s12938-021-00863-x]

36 **La Salvia M**, Secco G, Torti E, Florimbi G, Guido L, Lago P, Salinaro F, Perlini S, Leporati F. Deep learning and lung ultrasound for Covid-19 pneumonia detection and severity classification. *Comput Biol Med* 2021; **136**: 104742 [PMID: 34388462 DOI: 10.1016/j.compbiomed.2021.104742]

37 **Mento F**, Perrone T, Fiengo A, Smargiassi A, Inchingolo R, Soldati G, Demi L. Deep learning applied to lung ultrasound videos for scoring COVID-19 patients: A multicenter study. *J Acoust Soc Am* 2021; **149**: 3626 [PMID: 34241100 DOI: 10.1121/10.0004855]

38 **Roy S**, Menapace W, Oei S, Luijten B, Fini E, Saltori C, Huijben I, Chennakeshava N, Mento F, Sentelli A, Peschiera E, Trevisan R, Maschietto G, Torri E, Inchingolo R, Smargiassi A, Soldati G, Rota P, Passerini A, van Sloun RJG, Ricci E, Demi L. Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound. *IEEE Trans Med Imaging* 2020; **39**: 2676-2687 [PMID: 32406829 DOI: 10.1109/TMI.2020.2994459]

39 **Sadik F**, Dastider AG, Fattah SA. SpecMEn-DL: spectral mask enhancement with deep learning models to predict COVID-19 from lung ultrasound videos. *Health Inf Sci Syst* 2021; **9**: 28 [PMID: 34257953 DOI: 10.1007/s13755-021-00154-8]

40 **Muhammad G**, Shamim Hossain M. COVID-19 and Non-COVID-19 Classification using Multi-layers Fusion From Lung Ultrasound Images. *Inf Fusion* 2021; **72**: 80-88 [PMID: 33649704 DOI: 10.1016/j.inffus.2021.02.013]

41 **Tsai CH**, van der Burgt J, Vukovic D, Kaur N, Demi L, Canty D, Wang A, Royse A, Royse C, Haji K, Dowling J, Chetty G, Fontanarosa D. Automatic deep learning-based pleural effusion classification in lung ultrasound images for respiratory pathology diagnosis. *Phys Med* 2021; **83**: 38-45 [PMID: 33706149 DOI: 10.1016/j.ejmp.2021.02.023]

42 **Xue W**, Cao C, Liu J, Duan Y, Cao H, Wang J, Tao X, Chen Z, Wu M, Zhang J, Sun H, Jin Y, Yang X, Huang R, Xiang F, Song Y, You M, Zhang W, Jiang L, Zhang Z, Kong S, Tian Y, Zhang L, Ni D, Xie M. Modality alignment contrastive learning for severity assessment of COVID-19 from lung ultrasound and clinical information. *Med Image Anal* 2021; **69**: 101975 [PMID: 33550007 DOI: 10.1016/j.media.2021.101975]

43 **Selvaraju RR**, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV); 2017 Oct 22-29; Venice, Italy. IEEE, 2017: 618–626 [DOI: 10.1109/ICCV.2017.74]

44 **Soldati G**, Smargiassi A, Inchingolo R, Buonsenso D, Perrone T, Briganti DF, Perlini S, Torri E, Mariani A, Mossolani EE, Tursi F, Mento F, Demi L. Proposal for International Standardization of the Use of Lung Ultrasound for Patients With COVID-19: A Simple, Quantitative, Reproducible Method. *J Ultrasound Med* 2020; **39**: 1413-1419 [PMID: 32227492 DOI: 10.1002/jum.15285]

45 **Howard AG**, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam M. Mobilenets: Efficient convolutional neural networks for mobile vision applications; 2017. Preprint. Cited 17 April 2017

46 **Zhou Z**, Sodha V, Rahman Siddiquee MM, Feng R, Tajbakhsh N, Gotway MB, Liang J. Models genesis: Generic autodidactic models for 3D medical image analysis. In Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap PT, Khan A editors. Medical Image Computing and Computer Assisted Intervention–MICCAI 2019. Proceedings of 22nd International Conference MICCAI; 2019 Oct 13-17; Shenzhen, China. Lecture Notes in Computer Science. Springer, Cham, Switzerland, 2019; 11767: 384-393 [DOI: 10.1007/978-3-030-32251-9_42]

47 **Krizhevsky A**, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ. Advances in neural information processing systems. New York: Curran Associates Inc, 2012; 25: 1097–1105

48 **Russakovsky O**, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, et al ImageNet large scale visual recognition challenge. Int J Comput Vis 2015; 115: 211–252 [DOI: 10.1007/s11263-015-0816-y]

49 **Horry MJ**, Chakraborty S, Paul M, Ulhaq A, Pradhan B, Saha M, Shukla N. Covid-19 detection through transfer learning using multimodal imaging data. IEEE Access 2020; 8: 149808–149824 [DOI: 10.1109/ACCESS.2020.3016780]

50 **Moran M**, Faria M, Giraldi G, Bastos L, Oliveira L, Conci A. Classification of Approximal Caries in Bitewing Radiographs Using Convolutional Neural Networks. *Sensors (Basel)* 2021; **21** [PMID: 34372429 DOI: 10.3390/s21155192]

51 **LeCun Y**, Haffner P, Bottou L, Bengio Y. Object recognition with gradient-based learning. In: Shape, Contour and Grouping in Computer Vision. Lecture Notes in Computer Science, vol 1681. Springer: Berlin/Heidelberg, Germany, 1999 [DOI: 10.1007/3-540-46805-6_19]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com