Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success

SA-CME

James H. Thrall, MD, Xiang Li, PhD, Quanzheng Li, PhD, Cinthia Cruz, MD, Synho Do, PhD, Keith Dreyer, DO, PhD, James Brink, MD

Credits awarded for this enduring activity are designated "SA-CME" by the American Board of Radiology (ABR) and qualify toward fulfilling requirements for Maintenance of Certification (MOC) Part II: Lifelong Learning and Self-assessment. To access the SA-CME activity visit https://3s.acr.org/Presenters/CaseScript/CaseView?CDId=0GSf930sE7s%3d.

Abstract

Worldwide interest in artificial intelligence (AI) applications, including imaging, is high and growing rapidly, fueled by availability of large datasets ("big data"), substantial advances in computing power, and new deep-learning algorithms. Apart from developing new AI methods per se, there are many opportunities and challenges for the imaging community, including the development of a common nomenclature, better ways to share image data, and standards for validating AI program use across different imaging platforms and patient populations. AI surveillance programs may help radiologists prioritize work lists by identifying suspicious or positive cases for early review. AI programs can be used to extract "radiomic" information from images not discernible by visual inspection, potentially increasing the diagnostic and prognostic value derived from image datasets. Predictions have been made that suggest AI will put radiologists out of business. This issue has been overstated, and it is much more likely that radiologists will beneficially incorporate AI methods into their practices. Current limitations in availability of technical expertise and even computing power will be resolved over time and can also be addressed by remote access solutions. Success for AI in imaging will be measured by value created: increased diagnostic certainty, faster turnaround, better outcomes for patients, and better quality of work life for radiologists. AI offers a new and promising set of methods for analyzing image data. Radiologists will explore these new pathways and are likely to play a leading role in medical applications of AI.

Key Words: Artificial intelligence, machine learning, opportunities, challenges, pitfalls

J Am Coll Radiol 2018;15:504-508. Copyright © 2017 American College of Radiology

INTRODUCTION

The invention of the programmable digital computer in the 1940s stimulated mathematicians and philosophers to speculate on the limits of what machines could do. Could machines learn to think? How close could machine capabilities come to those of human beings? A conference at Dartmouth University in 1956 explored these questions and led to the coining of the term *artificial intelligence* (AI) [1]. The race was on.

In the ensuing 60 years, enthusiasm for AI has waxed and waned but has reignited recently with the availability of ever less expense, massively parallel computing systems. The term *deep learning* was added to the AI lexicon to reflect the ability to harness new computing power to develop more powerful AI approaches with more layers of analysis than heretofore possible. The successes of AI programs from IBM (Armonk, New York) in the games of chess and the quiz show *Jeopardy!* (Deep Blue) and from Google (Mountain View, California) in the game of *Go* (DeepMind) [2] were exciting milestones that made people outside of the scientific community aware of AI and its potential.

Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts.

Corresponding author and reprints: James H. Thrall, MD, Department of Radiology Massachusetts General Hospital, 25 New Chardon Street, #400, Boston, MA 02109; e-mail: jthrall@mgh.harvard.edu.

Dr Li has an institutional disclosure for the system: "Automatic pre-screening method for pneumothorax detection." Dr Thrall has an institutional disclosure on analysis of liver lesions. No patent has been applied for. The other authors have no conflicts of interest related to the material discussed in this article.

1546-1440/17/$36.00 ■ https://doi.org/10.1016/j.jacr.2017.12.026

Major corporations and governments around the world have embraced AI technology as one of the important strategies for dealing with the enormous amounts of digital data being generated in the information age—the age of "big data." AI is also on the doorstep of medical practice. The trickle of publications now appearing in journals will soon turn into a flood.

The radiology community has played a leading role in propelling medicine into its digital age and now has the opportunity to become a leader in exploring medical applications of AI. The tens of millions of radiology reports and billions of images now archived in digital form exemplify the concept of "big data" and constitute the required substrate for AI research.

The fundamental question is whether AI applications in radiology can add value. Adding value includes the discovery of new knowledge and extraction of more and better information from imaging examinations to achieve better outcomes for patients at lower cost. For radiologists, adding value includes establishment of more efficient work processes and improved job satisfaction.

The goal of this perspective is to help create a framework—apart from a discussion of AI technology per se—for developing strategies to explore the potential of AI in radiology and to identify a number of scientific, cultural, educational, and ethical issues that need to be addressed.

OPPORTUNITIES

Two areas of opportunity that can help provide a framework for approaching AI in imaging deserve discussion: the desirability of establishing standards and infrastructure and the opportunity to establish a categorical model for approaching the spectrum of clinical and research applications of AI to help identify and understand their respective value propositions.

Standards and Infrastructure

AI imaging research would benefit from the establishment of (1) national and international image sharing networks, (2) reference datasets of proven cases against which AI programs can be tested and compared, (3) criteria for standardization and optimization of imaging protocols for use in AI applications, and (4) a common lexicon for describing and reporting AI applications.

Access to large numbers of proven cases is necessary to test and validate AI programs and, for many applications, to train them. Taking face recognition as an example, the DeepFace system used 4.4 million labeled faces from 4,030 people collected by Facebook (Menlo Park, California) as training data. Its accuracy consistently approaches human-level performance [3]. Thus, for the purpose of analyzing medical images, even for relatively common conditions, it would be advantageous to be able to aggregate material from multiple institutions. The technical basis for accomplishing this already exists through networks created by the RSNA and the ACR as well as work such as The Cancer Imaging Archive at the National Cancer Institute to establish systems for image sharing [4]. The existence of these networks represents a unique opportunity to create shared image repositories for AI research. The new ACR Data Science Institute could serve as a convening organization.

As a corollary to the foregoing, robust methods are needed for quality control of shared images and ensuring the integrity of image data. Standards still need to be developed that address curation of images. If image data are corrupted in transmission or storage or during processing, it will be difficult to duplicate work or confirm its validity. Although several scientific informatics systems to support data transfer, storage, quality control, and query have been applied across institutions, such as those provided by the National Center for Biotechnology Information [4], the complete needs of such a shared informatics system are not yet defined for radiology.

An image sharing network would support and facilitate use of reference datasets of proven cases on which to test and compare new AI programs for accuracy and other measures of performance, such as required processing power. Such datasets also could be used in optimizing programs and selecting the best programs for further development and clinical use. Ideally, a sufficient number of reference datasets would be established in each application area to reflect the demographics of different patient populations. A database conceptually similar to the "ImageNet" [5], which has been used as the benchmark for object recognition tasks, would be valuable for advancing AI in radiology.

The high variability in imaging protocols between institutions and even variability in the execution of a given protocol within an institution are potential impediments to development and use of AI applications in imaging [6]. For example, subjective analysis of a CT scan may be somewhat tolerant of the timing of contrast material administration, and variable timing can create an "apples-to-oranges" problem for AI programs that rely on quantitative factors and may require scaling or a background subtraction step. The impact of protocol variability needs to be more fully studied.