

74297\_Auto\_Edited.docx

**Name of Journal:** *Artificial Intelligence in Medical Imaging*

**Manuscript NO:** 74297

**Manuscript Type:** MINIREVIEWS

## **Artificial intelligence: Advances and new frontiers in medical imaging**

Marc R Fromherz, Mina S Makary

### **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 (CAD), 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 PAC system 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.

### **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<sup>[1,2]</sup>. 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<sup>[1,2,3]</sup>. 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<sup>[3,4]</sup>. This demand, combined with declining reimbursement, has put more pressure on radiologists to increase productivity<sup>[5]</sup>. 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<sup>[3,6,7]</sup>.

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 FDA approved products in 2020, however the once foreseen AI takeover has not yet manifested <sup>[8,9,10]</sup>. 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<sup>[11,12]</sup>. 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”<sup>[13]</sup>.

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<sup>[3,4,14]</sup>. 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, 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<sup>[15]</sup>. 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 medical record system and compare it to established guidelines, the system can then make the appropriate recommendation<sup>[16,17,18]</sup>. With an estimated 10% of all imaging studies being ordered in error, these nonpixel-based algorithms can automatically detect and eliminate erroneous study orders<sup>[19,20]</sup>.

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<sup>[11]</sup>. 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<sup>[11]</sup>.

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<sup>[21]</sup>. 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<sup>[22,23]</sup>.

### ***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<sup>[14]</sup>. 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<sup>[24]</sup>. A study by Sanford *et al* 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<sup>[24]</sup>. 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<sup>[1,2,25]</sup>. 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<sup>[26]</sup>. 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%<sup>[26]</sup>. This is a representative example of new AI algorithms geared toward instantaneous, automatic interpretation.

**Table 1. Areas of radiology workflow with current AI implementation**

## **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 which have some specific limitations of their own<sup>[1,27]</sup>.

### ***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<sup>[3]</sup>. 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<sup>[28,29]</sup>. 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<sup>[28]</sup>. 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.

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

### *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<sup>[3,30,31]</sup>. 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<sup>[32]</sup>. 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<sup>[28,29]</sup>. If multiple different AI systems are needed for each specific pathology the process of creating and implementing these systems may not be fiscally feasible<sup>[33]</sup>. 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<sup>[8,34]</sup>. 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<sup>[8,34,35]</sup>. 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.

In the previously mentioned Shaffter *et al* 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<sup>[26]</sup>. 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<sup>[36]</sup>. 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<sup>[37]</sup>. 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<sup>[8,9,10]</sup>. 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<sup>[28,29]</sup>. 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<sup>[38]</sup>. 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 widespread 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 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<sup>[39,40]</sup>. 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 days<sup>[30]</sup>. Another study on triaging patients based on head CT findings produced similar results, with an AUC of 0.92 for accurately detecting intracranial hemorrhage<sup>[41]</sup>. 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<sup>[30,41]</sup>.

**Figure 2. Deep-learning AI application in triaging head CT images. The input image characteristics are extracted and analyzed by the CNN to create an output. The output is then flagged or not flagged depending on the algorithm’s interpretation.**

### *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<sup>[3]</sup>. One of the obvious applications for this is oncology, with AI models already demonstrating their ability to accurately measure therapeutic response and tumor recurrence<sup>[42,43]</sup>. Throughout the 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<sup>[44,45,46]</sup>. Not only does this present the potential to make educated decisions for COVID 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<sup>[47]</sup>. 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.

ORIGINALITY REPORT

0%

SIMILARITY INDEX

PRIMARY SOURCES

EXCLUDE QUOTES ON

EXCLUDE MATCHES OFF

EXCLUDE BIBLIOGRAPHY ON