Is Machine Learning All You Need for Radiology Analysis?

21104272

"We should stop training radiologists now. It's just completely obvious that within five years, deep learning is going to do better than radiologists" declared by Geoffrey Hinton, one of the Turing Award Winner 2018, in 2016. However, 6 years have passed, is machine learning powerful enough to replace the radiologists for radiology?

1 Introduction to Radiology

Radiology is a branch of medical science that uses imaging technology and radiation to diagnose and treat diseases [LSB+12]. The origins of diagnostic radiology can be traced back to 1895, at the time that Wilhelm Röntgen made the first X-ray film image of his wife's hand [Rön96], as shown in 1. It is a dynamic discipline, and with the continuous improvement of technologies such as physics, electrical engineering and computer science, it has changed rapidly: not only the number of imaging methods has increased, but each method is constantly improving and perfecting its application in medical diagnosis [BH12].

Figure 1: The first X-ray photograph named "Hand with Rings", display the left hand of Wilhelm Röntgen's wife, Anna Bertha Ludwig (1895)

Major imaging methods of radiology can be considered as Conventional Radiography which including X-ray radiography [YR97]; Cross-Sectional Techniques, such as Computed Tomography (CT), Magnetic Resonance (MR), and Ultrasonography (US) [PDBT01]; Radiographic Contrast Agents that contains Iodinated Contrast Agents, Intravascular Contrast Agents and Gastrointestinal Contrast Agents [Bet87]; and Nuclear Radiology [MJHYM08].

However, after obtaining medical images by radiology techniques, it still needs to be analyzed and diagnosed by people who have professional knowledge, called radiologist. Radiologist is usually a well-trained doctor that responsible for diagnosing and treating injury and disease for patients with radiography equipment and techniques [LSB+12]. In United States of America, a radiologist need to have a Doctor of Medicine (M.D.) or Doctor of Osteopathic Medicine (D.O.) degree, and based on American Board of Radiology [oR], it's usually take at least 13 years averagely to become a radiologist after entering college, including 4 years for undergraduate degree, 4 years medical school, 1 or 2 years internship, and then 4 years of residency training in diagnostic radiology.

There's no doubt that under such a high threshold, the number of radiologist is very rare. From the statistics of Statista [Elf21], in 2021 there are more than 48,000 active radiology physicians in the U.S., which is around 0.015% of the total population in the country. With such scarcity, it can be expected that the cost of seeing a radiologist is expensive, such as registration fees, diagnosis and treatment fees, and time required waiting for the appointment. In addition, as the great improvement of radiology equipment these days, much more higher resolution medical images can be acquired and with that comes the increase of radiologists' workload [WS12]. So, whether there exist an alternative way which can be more rapid and convenient to help people analyze their medical images?

2 Machine Learning in Radiology

Machine learning, a sub-field of Artificial Intelligence (AI), studies and constructs a special algorithm rather than a specific algorithm that allows the computer to learn from the data to make predictions [Sam67]. It can use the gathered data to build an appropriate model automatically, and then use such model to predict new situations. Applying in the field of medicine, it not only can used to detect non-obvious or even hidden correlations and relationships between patients' data, but also for research and for the clinical practice [CJ20]. Machine Learning, particularly deep learning, has demonstrated extraordinary progress in the field of image processing and become a remarkably strengthful tool for image-recognition tasks [KNH+21]. With collected large clinical datasets, AI is becoming a major component of many applications in healthcare, including ophthalmology [GPC+16], dermatology [EKN+17], drug discovery [GHS16].

Back to radiology, as mentioned before, radiological imaging data continues to grow at a much higher rate than the number of radiologists [BGM09]. Based on [MSE+15], a radiologist needs to explain an image every 3-4 seconds over an 8-hour working day averagely, and under such high workloads and restricted conditions further errors could be unavoidably happened. As Andrew Ng, a pioneer of deep learning said: "I want to live in an AI-powered society. When anyone goes to see a doctor, I want AI to help that doctor provide higher quality and lower cost medical service." The main driving force for the application of AI in radiology imaging is the desire for more efficient and effective healthcare. Compare to diagnosis radiological image manually by radiologists, two advantages make deep learning increase care quality while lowering costs in radiology: with the state-of-art network architecture, deep learning can had higher sensitivities than radiologists for some classification tasks [WZL+17], while sensitivity can be understood as the proportion of people who really get disease who are judged to be disease [ACG+06]; and another is that, with sufficient data, deep learning could be developed and deployed rapidly without any domain knowledge due to deep learning's strength on extracting data representation automatically [BCV13]. [MBSB19] shows that with similar accuracy, data-driven AI is much faster than radiologists can on clinical images' inference.

Moreover, machine learning can also benefit current technical challenges in radiology, such as imaging segmentation, imaging classification, and disease detection [MAC+18]. Compare with human, features of deep learning have made them have more significant advantages. One of the strengths is that AI could detect subtle details that are difficult for the human eye to find. For some of the unusual details in the clinical image, even the best radiologists are often difficult to distinguish whether it's healthy or diseased [LYS+19]. Within deep learning, the model can not only diagnose the tiny details in the image but also provide more accurate risk prediction for breast cancer [YLS+19]. Also, AI can distinguishing similar CT characteristics between different pathema well, like COVID-19 and pneumonia of other diseases, and [BWX+20] indicates that model can achieve more than 87% in accuracy and sensitivity. Another character that AI beyond radiologists is that it can perform analytical tasks that exceed human capabilities due to its ability for handling large amount of data [NVK+15]. For example, although the segmentation of three-dimensional ultrasound volumes into functional tissues can greatly contribute to the accuracy of clinical diagnosis of breast cancer, manual segmentation of the entire 3D volume is not practical in clinical because it's time-consuming and resource-intensive [SBV+93]. Training with massive data, deep learning reveals a powerful capability on segmenting 3D radiology information into fibro-glandular tissue, mass, and fatty tissue, and shows a great compatibility with manual segmentation (85.7%), but with a much more rapidly speed [GLR+16]. Figure 2 gives an example of deep learning in application of radiology diagnose [HPQ+18].

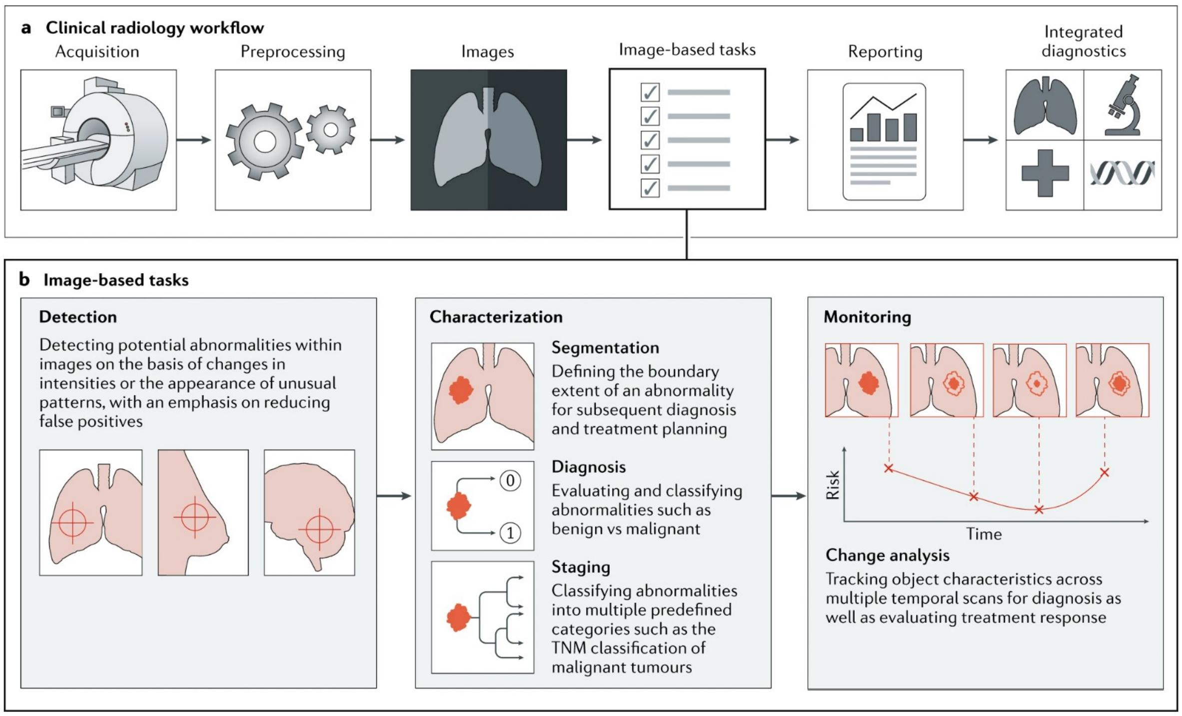

Figure 2: An example of deep learning in application of radiology diagnose

That's why Geoffrey Hinton, the 'godfather' of deep learning, who met and promoted the development of deep learning, claimed as early as 2016: "We should stop training radiologists now. It's just completely obvious that within five years, deep learning is going to do better than radiologists". In fact, this view was accepted by many people, and based on the the survey in 2018, even more than 68% interviewed medical students believed the demand of radiologists would be reduced by AI [Rea19]. Hinton's point of view is also proved by US Food and Drug Administration (FDA): they approved first AI algorithm for medical imaging the same year Hinton made his claim in 2016. Moreover, in 2020 there are more than 80 AI algorithms had been approved in the US [BDM20]. In both the US and Europe, 58% out of 129 devices in the USA and 53% out of 126 devices in Europe, AI/ML-based medical devices are approved for radiology by 2021 [MDV21].

These statistics seems to indicate the trend that machine learning is conquering medical area of radiology step by step, but unfortunately, this approach is hitting a wall.

3 Challenges for Radiology with AI

Instead of completely replacing radiologists, artificial intelligence still has a long way to go. In the United States, the number of radiologists has increased by 7% between 2015-2019 [DRJ+19]. Moreover, reports from more than one source still indicated the growing physician shortage of radiologists for current and for next decades [FCB21], and they predict that the position of radiologists would be continuously increased for following years [oAMC20]. In addition, according to the results from a survey by American College of Radiology (ACR), only around 30% out of 1,427 radiologists using AI for their regular clinical diagnostician and treatment, and the reason most of them don't AI for automatically clinical use is due to its inconsistent performance [AAC+21]. On the one hand, there is the increasing number of AI-based algorithm and equipment for medical to be invented and approved, and on the other hand, there is the shortage of radiologists and the radiologists' disapproval of AI. Where is the problem? Are there exist ways to solve?

3.1 Robustness

Indeed, challenges exist in deploying radiological deep learning solutions in clinical practice at current, but there do have possibilities for improvement by technical methods.

One of the challenges is that the input mode is single for the existing model. Not only through images, radiologists can have a better understanding of the disease from multiple aspects, such as the patient's description of their own symptoms, so that they can make more accurate judgments [PBHK16]. Multi-modal machine learning (MMML) have already shown its ability for fusion with multiple modals, like image and text, image and audio, to make a better performance than single-modal model on prediction tasks in recent years [BAM18]. By inputting images and descriptions of the disease, a multi-modal deep learning model can also have the ability to to diagnose diseases by combining various type of information like a radiologist [CYW+22]. In addition, MMML can also help with the challenge of the interpretability for deep learning. As widely known deep learning is a black-box model, and due to the lack of interpretability, doctors were difficult to trust it when the model diagnosis results are different from themselves [ZZ18]. Recent works about MMML in radiology offered a potential improvement probability. Based on the translation technique of MMML, the model can make a interpretation of input medical images by providing text of disease description as output, and doctors can have more information about AI-generated predictions, rather than a binary classification result only, based on these descriptions [NMAdS+19]. Another challenge is that there is currently no AI system which is capable to detect comprehensive abnormalities in the entire body like radiologists [SBK+19]. By recognizing multiple different sections, segmenting them and then making a classification in the given medical image, multi-task learning approach might be a solution for this problem [ZY21]. Another problem that affects model performance is the ability of real-world generalization outside of the training set [ZBL+18]. Besides collecting with more data, regularization approaches such as dropout, batch normalization and data augmentation, and model optimization could be helpful against such challenge [NBMS17].

3.2 Fairness

Studies have shown that human biases, such as gender and race, can be amplified by AI in a variety of situations [ZS18][H+17], and this would reproduce social stereotypes of minorities on the model, resulting in poor performance of the model, which is particularly dangerous in healthcare [CSM18]. Therefore, the choice of dataset and their features, especially in medical images, is important since it would affect the trained models and so that the conclusions would generate bias when making predictions [VC21].

While obtaining diverse and balanced raw medical data is a challenge, their collection often relies on publicly available datasets, such as anonymous medical images that patients agree to use for research or product development. Unfortunately, these datasets do not always represent the same demographic data [LAD+19]. When data collection becomes a problem, solve the problem of data imbalance technically could be considered, and this normally has two perspectives: from data and from model [XWY19]. When the solution of imbalanced data be treated as data level, sampling method such as data augmentation can be applied for minority labels to change to distribution of datasets, and therefore made data become balance during the training process of AI model [SK19]. For algorithmic levels, modifying the deep learning model approaches like cost-sensitive learning and threshold moving are proved to be applicable for making prediction on imbalanced dataset [JK19].

3.3 Ethics

Ethics is also an important point that can not be ignored when deploying deep learning on radiology data. Well-organized medical data is hard to obtain, not only because the considerable time and expense to be taken but also due to its highly sensitive and tightly regulated usuage [RHL+20]. In addition, research indicated that even delete the metadata, such as name and birthday, patients still can be re-identified, so it's still not enough to protect patient's privacy [RHDM19]. This cause a further reluctance for sharing medical data since once the healthcare data is leaked or sold by private companies, patients may be at high risk of discrimination in areas such as insurance and employment, resulting in the inability to have a normal life [Mei02].

Technically, training AI models with non-shared input datasets and then merging them as a shared model can be used as a way to protect data privacy [PAH+17]. Besides, federal learning, which is a decentralized learning paradigm, can solve data governance and privacy issues through collaborative training algorithms that do not exchange data itself [YLCT19]. By keeping the radiology data localized during training, the shared model is learned by federated learning with local updates, and its inference is performed locally on a real-time copy of the shared model so that to eliminate possible privacy breaches from healthcare data sharing [LMX+19].

4 Future Perspectives

In 2022, we don't see any radiologist is replaced by AI. Instead, the application of machine learning in radiology is more difficult than it seems. Critics have even suggested that "Few fields have been more filled with hype and bravado than artificial intelligence. It has flitted from fad to fad decade by decade, always promising the moon, and only occasionally delivering" [Mar]. However, never in human history has there been a pattern recognition model as powerful as deep learning [HSW89], and its potential is huge enough to worth further research. For now, we still need to keep training radiologists. However, AI offers a possibility to increase quality of care while lowering costs in radiology, and with the improvement of AI these years, in field such as multi-modal machine learning, meta learning, and federal learning, we have reason to expect that AI can change its status as a tool to assist radiologists [CS19] and become a real reliable diagnostic tool one day.

References

[AAC+21] Bibb Allen, Sheela Agarwal, Laura Coombs, Christoph Wald, and Keith Dreyer. 2020 acr data science institute artificial intelligence survey. *Journal of the American College of Radiology*, 18(8):1153–1159, 2021.

[ACG+06] Silvio Aime, S Geninatti Crich, Eliana Gianolio, Giovanni Battista Giovenzana, Lorenzo Tei, and Enzo Terreno. High sensitivity lanthanide (iii) based probes for mr-medical imaging. *Coordination Chemistry Reviews*, 250(11-12):1562–1579, 2006.

[BAM18] Tadas Baltrušaitis, Chaitanya Ahuja, and Louis-Philippe Morency. Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, 41(2):423–443, 2018.

[BCV13] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.

[BDM20] Stan Benjamens, Pranavsingh Dhunnoo, and Bertalan Meskó. The state of artificial intelligence-based fda-approved medical devices and algorithms: an online database. *NPJ digital medicine*, 3(1):1–8, 2020.

[Bet87] Michael A Bettmann. Radiographic contrast agents—a perspective, 1987.

[BGM09] Giles WL Boland, Alexander S Guimaraes, and Peter R Mueller. The radiologist's conundrum: benefits and costs of increasing ct capacity and utilization. *European radiology*, 19(1):9–11, 2009.

[BH12] William E Brant and Clyde A Helms. Fundamentals of diagnostic radiology. 2012.

[BWX+20] Harrison X Bai, Robin Wang, Zeng Xiong, Ben Hsieh, Ken Chang, Kasey Halsey, Thi My Linh Tran, Ji Whae Choi, Dong-Cui Wang, Lin-Bo Shi, et al. Artificial intelligence augmentation of radiologist performance in distinguishing covid-19 from pneumonia of other origin at chest ct. *Radiology*, 296(3):E156–E165, 2020.

[CJ20] Davide Chicco and Giuseppe Jurman. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC medical informatics and decision making*, 20(1):1–16, 2020.

[CS19] Stephen Chan and Eliot L Siegel. Will machine learning end the viability of radiology as a thriving medical specialty? *The British journal of radiology*, 92(1094):20180416, 2019.

[CSM18] Danton S Char, Nigam H Shah, and David Magnus. Implementing machine learning in health care—addressing ethical challenges. *The New England journal of medicine*, 378(11):981, 2018.

[CYW+22] Can Cui, Haichun Yang, Yaohong Wang, Shilin Zhao, Zuhayr Asad, Lori A Coburn, Keith T Wilson, Bennett A Landman, and Yuankai Huo. Deep multi-modal fusion of image and non-image data in disease diagnosis and prognosis: A review. *arXiv preprint arXiv:2203.15588*, 2022.

[DRJ+19] T Dall, Ryan Reynolds, Kari Jones, Ritashree Chakrabarti, and W Iacobucci. The complexities of physician supply and demand: projections from 2017 to 2032. *IHS Markit Ltd for the Association of American Medical Colleges, Washington, DC*, 2019.

[EKN+17] Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639):115–118, 2017.

[Elf21] J Elflein. Number of active physicians in the us 2021 by specialty area, 2021.

[FCB21] Shannon G Farmakis, Jocelyn D Chertoff, and Richard A Barth. Pediatric radiologist workforce shortage: Action steps to resolve. *Journal of the American College of Radiology*, 18(12):1675–1677, 2021.

[GHS16] Erik Gawehn, Jan A Hiss, and Gisbert Schneider. Deep learning in drug discovery. *Molecular informatics*, 35(1):3–14, 2016.

[GLR+16] Peng Gu, Won-Mean Lee, Marilyn A Roubidoux, Jie Yuan, Xueding Wang, and Paul L Carson. Automated 3d ultrasound image segmentation to aid breast cancer image interpretation. *Ultrasonics*, 65:51–58, 2016.

[GPC+16] Varun Gulshan, Lily Peng, Marc Coram, Martin C Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tom Madams, Jorge Cuadros, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 316(22):2402–2410, 2016.

[H+17] Matthew Hutson et al. Even artificial intelligence can acquire biases against race and gender. *Science*, 10, 2017.

[HPQ+18] Ahmed Hosny, Chintan Parmar, John Quackenbush, Lawrence H Schwartz, and Hugo JWL Aerts. Artificial intelligence in radiology. *Nature Reviews Cancer*, 18(8):500–510, 2018.

[HSW89] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.

[JK19] Justin M Johnson and Taghi M Khoshgoftaar. Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1):1–54, 2019.

[KNH+21] Salman Khan, Muzammal Naseer, Munawar Hayat, Syed Waqas Zamir, Fahad Shahbaz Khan, and Mubarak Shah. Transformers in vision: A survey. *ACM Computing Surveys (CSUR)*, 2021.

[LAD+19] Jonathan M Loree, Seerat Anand, Arvind Dasari, Joseph M Unger, Anirudh Gothwal, Lee M Ellis, Gauri Varadhachary, Scott Kopetz, Michael J Overman, and Kanwal Raghav. Disparity of race reporting and representation in clinical trials leading to cancer drug approvals from 2008 to 2018. *JAMA oncology*, 5(10):e191870–e191870, 2019.

[LMX+19] Wenqi Li, Fausto Milletarì, Daguang Xu, Nicola Rieke, Jonny Hancox, Wentao Zhu, Maximilian Baust, Yan Cheng, Sébastien Ourselin, M Jorge Cardoso, et al. Privacy-preserving federated brain tumour segmentation. In *International workshop on machine learning in medical imaging*, pages 133–141. Springer, 2019.

[LSB+12] Lawrence Liebscher, Cynthia Sherry, Jonathan Breslau, Gerald Dodd, Howard Fleishon, Paul Larson, Carolyn Meltzer, and Richard Strax. The general radiologist in the 21st century. *Journal of the American College of Radiology*, 9(8):554–559, 2012.

[LYS+19] Constance D Lehman, Adam Yala, Tal Schuster, Brian Dontchos, Manisha Bahl, Kyle Swanson, and Regina Barzilay. Mammographic breast density assessment using deep learning: clinical implementation. *Radiology*, 290(1):52–58, 2019.

[MAC+18] Morgan P McBee, Omer A Awan, Andrew T Colucci, Comeron W Ghobadi, Nadja Kadom, Akash P Kansagra, Srini Tridandapani, and William F Auffermann. Deep learning in radiology. *Academic radiology*, 25(11):1472–1480, 2018.

[Mar] Gary Marcus. Deep learning is hitting a wall.

[MBSB19] Maciej A Mazurowski, Mateusz Buda, Ashirbani Saha, and Mustafa R Bashir. Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on mri. *Journal of magnetic resonance imaging*, 49(4):939–954, 2019.

[MDV21] Urs J Muehlematter, Paola Daniore, and Kerstin N Vokinger. Approval of artificial intelligence and machine learning-based medical devices in the usa and europe (2015–20): a comparative analysis. *The Lancet Digital Health*, 3(3):e195–e203, 2021.

[Mei02] Eileen Meier. Medical privacy and its value for patients. In *Seminars in oncology nursing*, volume 18, pages 105–108. Elsevier, 2002.

[MJHYM08] Fred A Mettler Jr, Walter Huda, Terry T Yoshizumi, and Mahadevappa Mahesh. Effective doses in radiology and diagnostic nuclear medicine: a catalog. *Radiology*, 248(1):254–263, 2008.

[MSE+15] Robert J McDonald, Kara M Schwartz, Laurence J Eckel, Felix E Diehn, Christopher H Hunt, Brian J Bartholmai, Bradley J Erickson, and David F Kallmes. The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. *Academic radiology*, 22(9):1191–1198, 2015.

[NBMS17] Behnam Neyshabur, Srinadh Bhojanapalli, David McAllester, and Nati Srebro. Exploring generalization in deep learning. *Advances in neural information processing systems*, 30, 2017.

[NMAdS+19] Nelson Nunes, Bruno Martins, Nuno André da Silva, Francisca Leite, and Mário J Silva. A multi-modal deep learning method for classifying chest radiology exams. In *EPIA Conference on Artificial Intelligence*, pages 323–335. Springer, 2019.

[NVK+15] Maryam M Najafabadi, Flavio Villanustre, Taghi M Khoshgoftaar, Naeem Seliya, Randall Wald, and Edin Muharemagic. Deep learning applications and challenges in big data analytics. *Journal of big data*, 2(1):1–21, 2015.

[oAMC20] Association of American Medical Colleges. The complexities of physician supply and demand: projections from 2018 to 2033, 2020.

[oR] American Board of Radiology. Diagnostic radiology. https://www.theabr.org/diagnostic-radiology.

[PAH+17] Le Trieu Phong, Yoshinori Aono, Takuya Hayashi, Lihua Wang, and Shiho Moriai. Privacy-preserving deep learning: Revisited and enhanced. In *International Conference on Applications and Techniques in Information Security*, pages 100–110. Springer, 2017.

[PBHK16] Ewoud Pons, Loes MM Braun, MG Myriam Hunink, and Jan A Kors. Natural language processing in radiology: a systematic review. *Radiology*, 279(2):329–343, 2016.

[PDBT01] Dominique Penninck, Gregory B Daniel, Robert Brawer, and Amy S Tidwell. Cross-sectional imaging techniques in veterinary ophthalmology. *Clinical techniques in small animal practice*, 16(1):22–39, 2001.

[Rea19] Sara Reardon. Rise of robot radiologists. *Nature*, 576(7787):S54–S54, 2019.

[RHDM19] Luc Rocher, Julien M Hendrickx, and Yves-Alexandre De Montjoye. Estimating the success of re-identifications in incomplete datasets using generative models. *Nature communications*, 10(1):1–9, 2019.

[RHL+20] Nicola Rieke, Jonny Hancox, Wenqi Li, Fausto Milletari, Holger R Roth, Shadi Albarqouni, Spyridon Bakas, Mathieu N Galtier, Bennett A Landman, Klaus Maier-Hein, et al. The future of digital health with federated learning. *NPJ digital medicine*, 3(1):1–7, 2020.

[Rön96] Wilhelm Conrad Röntgen. On a new kind of rays. *Science*, 3(59):227–231, 1896.

[Sam67] Arthur L Samuel. Some studies in machine learning using the game of checkers. ii—recent progress. *IBM Journal of research and development*, 11(6):601–617, 1967.

[SBK+19] Luca Saba, Mainak Biswas, Venkatanareshbabu Kuppili, Elisa Cuadrado Godia, Harman S Suri, Damodar Reddy Edla, Tomaž Omerzu, John R Laird, Narendra N Khanna, Sophie Mavrogeni, et al. The present and future of deep learning in radiology. *European journal of radiology*, 114:14–24, 2019.

[SBV+93] Paul Suetens, Erwin Bellon, Dirk Vandermeulen, M Smet, Guy Marchal, Johan Nuyts, and Luc Mortelmans. Image segmentation: methods and applications in diagnostic radiology and nuclear medicine. *European journal of radiology*, 17(1):14–21, 1993.

[SK19] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.

[VC21] Gaël Varoquaux and Veronika Cheplygina. How i failed machine learning in medical imaging–shortcomings and recommendations. *arXiv preprint arXiv:2103.10292*, 2021.

[WS12] Shijun Wang and Ronald M Summers. Machine learning and radiology. *Medical image analysis*, 16(5):933–951, 2012.

[WZL+17] Hongkai Wang, Zongwei Zhou, Yingci Li, Zhonghua Chen, Peiou Lu, Wenzhi Wang, Wanyu Liu, and Lijuan Yu. Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18f-fdg pet/ct images. *EJNMMI research*, 7(1):1–11, 2017.

[XWY19] XU Xiaolong, CHEN Wen, and SUN Yanfei. Over-sampling algorithm for imbalanced data classification. *Journal of Systems Engineering and Electronics*, 30(6):1182–1191, 2019.

[YLCT19] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2):1–19, 2019.

[YLS+19] Adam Yala, Constance Lehman, Tal Schuster, Tally Portnoi, and Regina Barzilay. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology*, 292(1):60–66, 2019.

[YR97] MJ Yaffe and JA Rowlands. X-ray detectors for digital radiography. *Physics in medicine & biology*, 42(1):1, 1997.

[ZBL+18] John R Zech, Marcus A Badgeley, Manway Liu, Anthony B Costa, Joseph J Titano, and Eric K Oermann. Confounding variables can degrade generalization performance of radiological deep learning models. *arXiv preprint arXiv:1807.00431*, 2018.

[ZS18] James Zou and Londa Schiebinger. Ai can be sexist and racist—it's time to make it fair, 2018.

[ZY21] Yu Zhang and Qiang Yang. A survey on multi-task learning. *IEEE Transactions on Knowledge and Data Engineering*, 2021.

[ZZ18] Quan-shi Zhang and Song-Chun Zhu. Visual interpretability for deep learning: a survey. *Frontiers of Information Technology & Electronic Engineering*, 19(1):27–39, 2018.