Prime Time for Artificial Intelligence in Interventional Radiology

Jarrel Seah[1,2] · Tom Boeken[3] · Marc Sapoval[3] · Gerard S. Goh[1,4,5]

Abstract Machine learning techniques, also known as artificial intelligence (AI), is about to dramatically change workflow and diagnostic capabilities in diagnostic radiology. The interest in AI in Interventional Radiology is rapidly gathering pace. With this early interest in AI in procedural medicine, IR could lead the way to AI research and clinical applications for all interventional medical fields. This review will address an overview of machine learning, radiomics and AI in the field of interventional radiology, enumerating the possible applications of such techniques, while also describing techniques to overcome the challenge of limited data when applying these techniques in interventional radiology. Lastly, this review will address common errors in research in this field and suggest pathways for those interested in learning and becoming involved about AI.

Keywords Machine learning · Interventional radiology · Deep learning · Artificial intelligence · AI

✉ Gerard S. Goh
g.goh@alfred.org.au

1   Department of Radiology, Alfred Health, Melbourne, VIC, Australia

2   Department of Neuroscience, Monash University, Melbourne, VIC, Australia

3   Vascular and Oncological Interventional Radiology, University of Paris, Hopital Européen Georges Pompidou, Paris, France

4   Department of Surgery, Central Clinical School, Monash University, Melbourne, VIC, Australia

5   National Trauma Research Institute, Central Clinical School, Monash University, Melbourne, VIC, Australia

Introduction

Artificial intelligence (AI) has been prominent in different fields including diagnostic radiology. AI breakthroughs are now also empowering the field of interventional radiology (IR) and recent surge in popularity was initially driven by the phenomenal success of deep neural networks in processing unstructured data such as images and audio through pattern recognition. The term AI has come to encompass all forms of machine learning (ML). In this review the term AI is used synonymously with ML, referring to techniques that construct predictive models from data, including deep learning, radiomics and other traditional machine learning techniques. The importance of the usage of AI is summarized in (Table 1).

A brief history of the field is difficult to summarize given its fragmented nature; however, the most popular techniques currently, artificial neural networks, date back to work by Rosenblatt[1] on the concept of the perceptron in the 1960s.

Artificial neural networks are inspired by the connectionist design of biological neural networks[2], comprising of artificial "neurons" which receive input from other neurons or the environment, performing a nonlinear activation function[3] on the sum of this input, and passing its output to other neurons. In short, each layer in an artificial neural network is a mathematical model loosely mimicking biological neurons, receiving information from one or more sources, processing the information and producing a response. This information is passed to other neurons for further analysis. To train a neural network, a large dataset of paired inputs and desired outputs, referred to as labels, must be collected. The training process then refers to altering the parameters (typically the weights) for each neuron such that the network produces the desired outputs

Table 1 The importance of artificial intelligence

Has the ability to incorporate and analyse a large amount of complex data rapidly

Identifies trends and patterns only partly detectable by humans
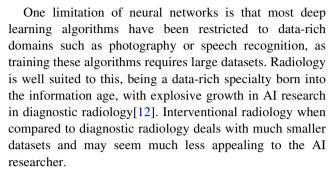
Independent of human bias in decision-making

Has the potential to aid interventional radiologists in the diagnosis, treatment and follow-up of disease

for each input. Currently, the most popular architecture for processing images is the convolutional neural network (CNN) [4], where neurons are organized spatially, taking inputs from adjacent pixels and processing it into higher-order representations before passing this information onto successive layers, eventually producing predictions. These predictions are compared to the labels, and the gradient of the parameters of the network, with respect to the error, is calculated through a process referred to as backpropagation[5]. Supervised learning is a type of ML that utilizes a set of input and output labelled training data [6] and can be used to estimate relationships between input and output parameters.

AI using artificial neural networks were recently re-popularized in 2012 by Krizhevsky et al. [7] with the development of large-scale parallel processing through graphics processing units combined with the availability of large datasets. In their revolutionary publication in 2012, Krizhevsky et al. [7] trained a deep CNN to classify over 1 million images (from the ImageNet contest) into 1000 different classes. Their network was built with 650,000 neurons and 60 million parameters. Though networks have since become even more complex, the AlexNet model created by Krizhevsky remains a reference for image classification. It is important to acknowledge that CNNs are still at the core of most research in this field at this nascent stage.

Recent research in medical imaging AI has focused on 3D CNNs such as 3D versions of popular 2D CNNs such as EfficientNet or Densenet. [8] Alternative techniques such as 2.5D networks (taking into account axial, coronal and sagittal planes) have also been developed. [9]. Even more recently in the wider AI space is the shift to purely attention-based mechanisms such as vision transformers [10] even for image-based tasks. These popular architectures are mentioned in brief as it is beyond the scope of this review article to describe these in detail. These approaches work remarkably well when networks are scaled up to millions or billions of parameters. Where traditional ML algorithms such as logistic regression tend to overfit with so many parameters, neural networks remarkably appear to work better with more parameters, particularly when multiple layers are stacked deeply. [11] Why this is the case remains one of the central mysteries of deep learning.

One limitation of neural networks is that most deep learning algorithms have been restricted to data-rich domains such as photography or speech recognition, as training these algorithms requires large datasets. Radiology is well suited to this, being a data-rich specialty born into the information age, with explosive growth in AI research in diagnostic radiology[12]. Interventional radiology when compared to diagnostic radiology deals with much smaller datasets and may seem much less appealing to the AI researcher.

In comparison with other interventional and procedural fields such as surgery or endoscopy, IR is data-rich. IR is one of a few specialties where a record is kept of the entire procedure in a standardized format, is available retrospectively and these datasets are mostly unexploited today. With the recent developments in few shot learning made by the deep learning community[13], novel techniques may drastically reduce the dataset size required for clinically effective algorithms, making it truly the prime time for AI in interventional radiology.

Applications

Potential applications of AI in IR can be divided into pre-procedural, peri-procedural and post procedural.

Pre-procedural Setting: Improving Patient Selection

Better patient selection is similar to the concept of precision medicine. AI decision support systems may help tailor treatment decisions based on imaging phenotypes, yielding better clinical results.

Interventional radiologists often rely on multidisciplinary boards for oncological treatment strategies. These board discussions perform multiparametric risk-stratification, integrating the patient's full data before a treatment is advised.

Several AI applications replicate and outperform these discussions by predicting the outcome from data available in each specialty (radiology, histology, molecular biology, etc.). The ability to incorporate clinical information, radiomics and genetic information may improve the objectivity and accuracy of decision-making. Such an

approach could potentially play a role in triaging patients for IR and subsequent therapy by assessing risks and making predictions about therapeutic outcomes [14]

An example of this application in the field of acute ischaemic stroke is the use of CT perfusion software in endovascular clot retrieval to estimate physiological parameters such as ischaemic core and penumbral volume, facilitating the selection of patients who are likely to have an optimal outcome. AI approaches to identifying patients for clot retrieval are being investigated, for instance based on CT angiography rather than CT perfusion. [15–17]

Pretreatment patient selection using deep learning is also being investigated for interventional oncology. Morshid et al. [18] describe an algorithm to predict response to transcatheter arterial chemoembolization for hepatocellular carcinoma (HCC) using pretreatment CT, combined with the clinical BCLC stage. They demonstrate that an AI model utilizing image and clinical features can outperform traditional staging systems in predicting benefit from TACE. Similarly, Peng et al.[19] developed a deep learning model that predicts response to TACE with an accuracy of 84%. Kim et al. [20] demonstrated that a combined radiomics and clinical model of HCC in response to TACE improved survival estimation when compared to clinical models alone. Other research in HCC treatment has found similar results when applied to surgical or thermo-ablative resection [21].

Multimodal planning may also integrate genetic information using AI models. Ziv et al. [22] trained a model to identify the genes most predictive of response to TACE and Kuo et al. [23] utilized radiomic analysis to identify imaging phenotypes associated with doxorubicin drug response gene expression in HCC.

Peri-procedural: Improving Procedures

AI can improve interventional procedures by accelerating computationally intensive or manual procedures, such as the correction of translational motion via pixel shifting in angiography. Traditional image registration techniques such as those proposed by Meijering et al. [24] are computationally intensive and have not had widespread uptake. Deep learning approaches may speed up corrected digital subtraction angiography, such as methods proposed by Gao et al. [25] which use generative adversarial networks to generate subtraction images without the preliminary non-contrast acquisition, avoiding the issue of translational motion entirely. This is achieved by acquiring a dataset of satisfactorily subtracted images paired with the unsubtracted images and training a neural network to predict the subtracted images from the unsubtracted image. This teaches the neural network anatomical and physical assumptions about the nature of angiographic contrast. The resultant neural network is capable of predicting the subtracted images from the unsubtracted angiographic images, without the use of the preliminary non-contrast acquisition.

Deep learning approaches have also been applied to identifying guidewire and catheters during angiography [26]. Such methods may permit more advanced algorithms such as virtual road mapping of the vasculature without contrast. Real-time AI registration algorithms could superimpose high-resolution preoperative imaging with procedural fluoroscopy, guiding the interventional radiologist during catheter manipulation.

AI-based ultrasound guidance [27] has been used in echocardiography to help guide the acquisition of echocardiograms. Deep learning algorithms estimate diagnostic quality of the image and suggest manoeuvres to improve the quality of such images. AI may provide recommendations on needle trajectory or other facets of interventional procedures, which may be particularly useful for novice operators.

The selection and personalization of endovascular devices is another area for AI. Yang et al. [28] used AI to segment and quantify stenosis on coronary angiography. Such algorithms could be used to objectively select the optimal stent for each lesion. Lee et al. [29] imagine a future where AI may guide the personalized 3D printing of cardiovascular stents.

Cho et al.[30] developed AI to predict fractional flow reserve of coronary lesions on angiography. This opens the way to extract hemodynamic parameters/physiological parameters from angiography, and AI may be able to even estimate flow distribution maps in the future.

AI has been proposed for skin dose estimation by taking into account angulation of the X-ray tube and tissue density. Radiation exposure during endoscopy has been reduced using an AI-equipped fluoroscopy unit with an ultrafast collimation system that reduced radiation exposure by $\sim$ 38%. Similar techniques could be used in interventional radiology. [31]

Respiratory motion compensation in PET/CT imaging has been implemented via elastic motion correction algorithms where AI determines a blurring kernel between a single motion corrected image and a single non-motion corrected target image. This results in a final image with reduced motion[32]. Similar applications could be applied in live fluoroscopy.

After Treatment: Improving Follow-Up

Following treatment, AI has a role to play in measuring response to treatment, prognostication and determining future management.

Most criteria used in diagnostic radiology for treatment response were not developed for interventional radiology,

which may lead to misevaluation during follow-up. AI research could help better assess these specific treatment responses. AI can be useful in longitudinal studies during follow-up of treatments to detect subtle changes between images identifying disease progress or recurrence earlier.

In oncology, automated volumetric measurements of tumour sizes or response evaluation criteria in solid tumours (RECIST) reads may be possible through deep learning. [33]. RECIST criteria themselves as a marker of response can be outperformed by AI. Dohan et al. [34] developed a radiomic signature that was able to predict overall survival and identify good responders better than RECIST1.1 criteria in patients with liver metastases from colorectal cancer treated with chemotherapy. The same models could be applied to IR treatment in liver metastases, potentially outperforming routine RECIST and equivalent criteria.

Procedural findings and histological features can also play a role in the choice of adjuvant therapy as suggested by Saillard et al.[35] who built a prediction model of survival after HCC resection based on pretherapeutic and histological preprocedural features.

Similarly, AI has a role in assessing response to treatment in acute ischaemic stroke. Thrombolysis in cerebral infarction (TICI) scores are often used to grade results following endovascular clot retrieval. AI can improve interobserver reliability and thereby improve the utility of such scores in prognosticating patients. [36, 37] AI algorithms may help reduce the time required to interpret post-treatment imaging and improve inter-observer variability. The ability of AI to extract quantitative metrics holds the promise of personalizing management plans, particularly in complex chronic conditions such as cancer.

While genetics and molecular pathology have played a large role in precision medicine, pre- and post-treatment imaging may identify additional disease phenotypes as well as quantify intervention success, which may help fine-tune management by prognosticating as well as determining the timing and need for follow-up imaging. [38]

Practical Challenges

The breadth of potential applications of AI in interventional radiology has seen a rise in academic papers published on this subject. Such projects face a common set of challenges..

The major challenge facing AI in interventional radiology is the relatively small dataset sizes when compared to diagnostic radiology, or in fact, to other non-medical applications of AI entirely. For instance, ImageNet, a widely used natural imagery database, contains over 14 million images. [39] In contrast, most medical applications have dataset sizes in the hundreds to thousands of unique samples. Therefore, standard deep learning models are difficult to train from scratch.

Perhaps the most simple method to reduce the number of samples required for a useful model is transfer learning, where models trained on different datasets might be used as a starting point, as information that these models might have learned from other datasets may be translated to this setting as well. [40]

Other approaches include the use of handcrafted features, an approach popular in the "traditional" computer vision literature in the early 2000s. These approaches, also known as radiomics when applied to imaging, reduce the number of required samples as the model does not have to learn the low-level features itself.

Another approach is to use data augmentation – standard transformations like affine transforms, adjusting brightness and contrast are useful, but novel augmentation techniques like MixUp [41] may help researchers get more out of their data. Medical imaging is often acquired quite differently from natural imagery, through techniques such as tomographic reconstruction. This offers the opportunity for different types of augmentations to introduce artefacts which are more typical in this setting, such as physics-based data augmentation [42]. Although this technique was found to be unsuccessful in previous work, it may be prove to be useful in more challenging datasets.

Recent developments in the deep learning field in semi-supervised learning, such as few-shot and zero-shot learning techniques [13, 43–46], may also help reduce the number of labelled samples required, by using unsupervised datasets as additional information.

For categorical data, such as models using clinical variables like age and sex, oversampling techniques such as synthetic minority oversampling technique (SMOTE) [47] may help generate synthetic data points that may improve an AI model's performance.

Small dataset sizes also exacerbate common mistakes made in AI projects. The use of checklists [48] may help prevent some of these avoidable errors. Common errors include the failure to split data by patient – i.e. including studies from a single patient in both the training and test datasets. This may lead to the model memorizing patient specific features, leading to over-optimistic results that do not translate into clinical practice. Other errors include not fully describing the hyper-parameter optimization process, or optimizing the hyper-parameters on the testing set, which again leads to over-estimation of the model's performance.

The training and testing datasets must also have defined inclusion and exclusion criteria to prevent "Frankenstein" datasets[48], where positive and negative cases are drawn from different sources, potentially leading to data leakage

as the AI model may recognize features specific to the dataset source rather than the disease of interest.

To prevent overstating the significance of any result, especially in small datasets, any measure of performance such as the accuracy or the area under the receiver operating characteristic (AUC) should be accompanied with confidence intervals. When applicable, AI models trained on images should be compared to a baseline clinical model using age, sex and other clinical features. If an AI model uses both imaging and clinical features, a sensitivity analysis should be performed by systematically modifying each input to assess its contribution on the final prediction. [49] In post-treatment and prognostication models, lack of complete follow-up in all participants is common and should be accounted for when measuring the accuracy of such models through censoring. [50]

Another source of unreliability stems from the constant evolution of clinical practice over time due to the introduction of new treatment approaches, technologies or changes in patient population [51]. Interventional radiology, in particular, is a rapidly evolving specialty with novel equipment and procedures constantly developing over time.

The use of AI in augmenting interventional radiologists is likely to increase as research in pretreatment, intratreatment and post-treatment applications translate into clinical practice. Due to the potential benefits and risks for patients, stringent prospective evaluation such as controlled trials should be undertaken where necessary to ensure that promising applications translate well.

How to Get Started in Ai

Given the promise of AI in interventional radiology many clinicians may wish to get involved in AI research and development. Key factors to be able to successfully translate a project into clinical practice include a clear understanding of the clinical benefits andadvantages of using AI, the availability of data measured in independent samples (typically at a patient level), the use of computing resources such as graphics processing units or tensor processing units and the technical skills to construct an AI model.

Specific steps around training and coding of AI models are beyond the scope of this review article; however, it is becoming easier the advent of open source deep learning framework libraries such as Pytorch [52] and Tensorflow [53]. A recommendation for interventional radiologists who are interested in learning more about AI is to begin with learning basic software and data carpentry skills in programming languages such as Python [54] and then expand knowledge by undertaking courses in frameworks such as Pytorch and Tensorflow.

Conclusion

The emergence of novel deep learning techniques and applications in interventional radiology is hugely exciting and offers multiple opportunities to aid in patient selection for intervention, improve patient care during interventional treatment and optimize post-treatment clinical follow-up. Interventional radiology with its smaller dataset sizes compared to diagnostic radiology stands to benefit from novel techniques such as semi-supervised learning, zero and few shot learning in the deep learning literature. The application of such techniques in interventional radiology must be rigorous and generalizable, and common errors must be avoided in order for successful clinical translation.

Declarations

Conflict of interest The other authors declare that they have no conflict of interest. Dr Jarrel Seah is a shareholder and consultant for Annalise.ai.

Consent for Publication For this type of study consent for publication is not required.

Ethical Approval All authors have reviewed and approved this manuscript.

References

1. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol Rev. 1958;65:386–408. https://doi.org/10.1037/h0042519.
2. Crick F. The recent excitement about neural networks. Nature. 1989;337:129–32.
3. Fukushima,. Visual Feature Extraction by a Multilayered Network of Analog Threshold Elements. IEEE Trans Syst Sci

Cybern. 1969;5:322–33. https://doi.org/10.1109/TSSC.1969. 300225.

4. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE. 1998;86:2278–323. https://doi.org/10.1109/5.726791.

5. Linnainmaa S. Taylor expansion of the accumulated rounding error. BIT. 1976;16:146–60. https://doi.org/10.1007/BF01931367.

6. B L, CJ W, J C,. The role of artificial intelligence in interventional oncology: a primer. J Vasc Interv Radiol. 2019;30:38-41.e1. https://doi.org/10.1016/J.JVIR.2018.08.032.

7. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Adv Neural Inform Process Syst. 2012;25:1097–105.

8. Liang S, Zhang R, Liang D, et al. Multimodal 3D DenseNet for IDH genotype prediction in gliomas. Genes. 2018. https://doi.org/10.3390/GENES9080382.

9. Yang J, Huang X, He Y, et al Reinventing 2D convolutions for 3D images. 2017;1

10. Dosovitskiy A, Beyer L, Kolesnikov A, et al An image is worth 16x16 words: transformers for image recognition at scale. 2020

11. Tan M, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks. 36th Int Conf Mach Learn ICML. 2019;10691–10700

12. Bluemke DA, Moy L, Bredella MA, et al. Assessing radiology research on artificial intelligence: a brief guide for authors, reviewers, and readers-from the radiology editorial board. Radiology. 2020;294:487–9.

13. Wang Y, Yao Q, Kwok JT, Ni LM. Generalizing from a few examples: a survey on few-shot learning. ACM Comput Surv. 2020;53:1–34. https://doi.org/10.1145/3386252.

14. Keshava SN, Kalva S. Artificial Intelligence in Interventional Radiology. J Clin Interv Radiol ISVIR. 2019;03:071–071. https://doi.org/10.1055/S-0039-1694309.

15. Hilbert A, Ramos LA, van Os HJA, et al. Data-efficient deep learning of radiological image data for outcome prediction after endovascular treatment of patients with acute ischemic stroke. Comput Biol Med. 2019;115: 103516. https://doi.org/10.1016/j.compbiomed.2019.103516.

16. Zeleňák K, Krajina A, Meyer L, Fiehler J, Intelligence EA, Behme D, Bulja D, Caroff J, Chotai AA, Da Ros V, Gentric JC. How to improve the management of acute ischemic stroke by modern technologies, artificial intelligence, and new treatment methods. Life. 2021;11(6):488.

17. Hofmeister J, Bernava G, Rosi A, et al. Clot-based radiomics predict a mechanical thrombectomy strategy for successful recanalization in acute ischemic stroke. Stroke. 2020;51:2488–94. https://doi.org/10.1161/STROKEAHA.120.030334.

18. Morshid A, Elsayes KM, Khalaf AM, et al. A machine learning model to predict hepatocellular carcinoma response to transcatheter arterial chemoembolization. Radiol Artif Intell. 2019;1: e180021. https://doi.org/10.1148/ryai.2019180021.

19. Peng J, Kang S, Ning Z, et al. Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from CT imaging. Eur Radiol. 2020;30:413–24. https://doi.org/10.1007/s00330-019-06318-1.

20. Kim J, Choi SJ, Lee SH, Lee HY, Park H. Predicting survival using pretreatment ct for patients with hepatocellular carcinoma treated with transarterial chemoembolization: comparison of models using radiomics. Am J Roentgenol. 2018;1:1026–34.

21. Ma X, Wei J, Gu D, et al. Preoperative radiomics nomogram for microvascular invasion prediction in hepatocellular carcinoma using contrast-enhanced CT. Eur Radiol. 2019;29:3595–605. https://doi.org/10.1007/s00330-018-5985-y.

22. Ziv E, Yarmohammadi H, Boas FE, et al. Gene signature associated with upregulation of the wnt/β-catenin signaling pathway predicts tumor response to transarterial embolization. J Vasc Interv Radiol. 2017;28:349-355.e1. https://doi.org/10.1016/j.jvir.2016.11.004.

23. Md K, J G, CB S, et al. Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. J Vasc Interv Radiol. 2007;18:821–30. https://doi.org/10.1016/J.JVIR.2007.04.031.

24. Meijering EHW, Zuiderveld KJ, Viergever MA. Image registration for digital subtraction angiography. Int J Comput Vis. 1999;31:227–46. https://doi.org/10.1023/a:1008074100927.

25. Gao Y, Song Y, Yin X, et al. Deep learning-based digital subtraction angiography image generation. Int J Comput Assist Radiol Surg. 2019;14:1775–84.

26. Ambrosini P, Ruijters D, Niessen WJ, et al. Fully Automatic and Real-Time Catheter Segmentation in X-Ray Fluoroscopy. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2017;10434:577–85.

27. Narang A, Bae R, Hong H, et al. Utility of a deep-learning algorithm to guide novices to acquire echocardiograms for limited diagnostic use. JAMA Cardiol. 2021. https://doi.org/10.1001/jamacardio.2021.0185.

28. Yang S, Kweon J, Roh JH, et al. Deep learning segmentation of major vessels in X-ray coronary angiography. Sci Rep. 2019;9:1–11. https://doi.org/10.1038/s41598-019-53254-7.

29. Lee Y, Veerubhotla K, Jeong MH, Lee CH. Deep learning in personalization of cardiovascular stents. J Cardiovasc Pharmacol Ther. 2020;25:110–20.

30. Cho H, Lee JG, Kang SJ, et al. Angiography-based machine learning for predicting fractional flow reserve in intermediate coronary artery lesions. J Am Heart Assoc. 2019. https://doi.org/10.1161/JAHA.118.011685.

31 Bang JY, Hough M, Hawes RH, Varadarajulu S. Use of artificial intelligence to reduce radiation exposure at fluoroscopy-guided endoscopic procedures. Am J Gastroenterol. 2020;115(4):555–61. https://doi.org/10.14309/ajg.0000000000000565.

32. Y L, K F, T M, et al. Respiratory motion compensation for pet/ct with motion information derived from matched attenuation-corrected gated PET data. J Nucl Med. 2018;59:1480–6. https://doi.org/10.2967/JNUMED.117.203000.

33 Tang Y, Harrison AP, Bagheri M, Xiao J, Summers RM. Semiautomatic recist labeling on ct scans with cascaded convolutional neural networks. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. Medical image computing and computer assisted intervention – MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part IV. Cham: Springer International Publishing; 2018. p. 405–13. https://doi.org/10.1007/978-3-030-00937-3_47.

34. Dohan A, Gallix B, Guiu B, et al. Early evaluation using a radiomic signature of unresectable hepatic metastases to predict outcome in patients with colorectal cancer treated with FOLFIRI and bevacizumab. Gut. 2020;69:531–9. https://doi.org/10.1136/gutjnl-2018-316407.

35. Saillard C, Schmauch B, Laifa O, et al. Predicting survival after hepatocellular carcinoma resection using deep learning on histological slides. Hepatology. 2020;72:2000–13. https://doi.org/10.1002/hep.31207.

36. Su R, Cornelissen SAP, van der Sluijs M, et al AutoTICI: Automatic brain tissue reperfusion scoring on 2D DSA images of acute ischemic stroke patients. 2020; arXiv XX:1

37. Nielsen M, Waldmann M, Frölich AM, et al. Deep learning-based automated thrombolysis in cerebral infarction scoring: a timely proof-of-principle study. Stroke. 2021;52:3497–504. https://doi.org/10.1161/STROKEAHA.120.033807.

38. Oakden-Rayner L, Carneiro G, Bessen T, et al. Precision Radiology: Predicting longevity using feature engineering and deep learning methods in a radiomics framework. Sci Rep. 2017;7:1–13. https://doi.org/10.1038/s41598-017-01931-w.

39. Deng J, Dong W, Socher R, et al. Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. 2009; pp. 248–255

40 Bozinovski S. Reminder of the first paper on transfer learning in neural networks, 1976. Informatica. 2020. https://doi.org/10.31449/inf.v44i3.2828.

41. Zhang H, Cisse M, Dauphin YN, Lopez-Paz D. mixup: Beyond Empirical Risk Minimization. 2017; arXiv

42. Omigbodun AO, Noo F, McNitt-Gray M, et al. The effects of physics-based data augmentation on the generalizability of deep neural networks: Demonstration on nodule false-positive reduction. Med Phys. 2019;46:4563–74. https://doi.org/10.1002/mp.13755.

43. Paul A, Shen TC, Lee S, et al. Generalized zero-shot chest X-ray diagnosis through trait-guided multi-view semantic embedding with self-training. IEEE Trans Med Imaging. 2021. https://doi.org/10.1109/TMI.2021.3054817.

44. Paul A, Tang YX, Shen TC, Summers RM. Discriminative ensemble learning for few-shot chest x-ray diagnosis. Med Image Anal. 2021;68: 101911. https://doi.org/10.1016/j.media.2020.101911.

45. Grill J-B, Strub F, Altché F, et al. Bootstrap your own latent: a new approach to self-supervised Learning. 2020; arXiv

46. Sohn K, Berthelot D, Li C-L, et al. FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence. 2020; arXiv

47. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321–57. https://doi.org/10.1613/jair.953.

48. Roberts M, Driggs D, Thorpe M, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell. 2021;3:199–217. https://doi.org/10.1038/s42256-021-00307-0.

49. Ankenbrand MJ, Shainberg L, Hock M, et al. Sensitivity analysis for interpretation of machine learning based segmentation models in cardiac MRI. BMC Med Imaging. 2021. https://doi.org/10.1186/s12880-021-00551-1.

50. Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. BMJ. 2020;369:26. https://doi.org/10.1136/bmj.m1328.

51. Avanzo M, Trianni A, Botta F, et al. Artificial intelligence and the medical physicist: welcome to the machine. Appl Sci. 2021. https://doi.org/10.3390/APP11041691.

52. Paszke A, Gross S, Massa F, et al. PyTorch: An imperative style, high-performance deep learning library. 2019

53. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. 2016

54 Wilson G. Software carpentry: lessons learned. F1000Res. 2016;3:62. https://doi.org/10.12688/f1000research.3-62.v2.