REVIEW ARTICLE

Next generation research applications for hybrid PET/MR and PET/CT imaging using deep learning

Greg Zaharchuk [1] (ID)

Abstract

Introduction Recently there have been significant advances in the field of machine learning and artificial intelligence (AI) centered around imaging-based applications such as computer vision. In particular, the tremendous power of deep learning algorithms, primarily based on convolutional neural network strategies, is becoming increasingly apparent and has already had direct impact on the fields of radiology and nuclear medicine. While most early applications of computer vision to radiological imaging have focused on classification of images into disease categories, it is also possible to use these methods to improve image quality. Hybrid imaging approaches, such as PET/MRI and PET/CT, are ideal for applying these methods.

Methods This review will give an overview of the application of AI to improve image quality for PET imaging directly and how the additional use of anatomic information from CT and MRI can lead to further benefits. For PET, these performance gains can be used to shorten imaging scan times, with improvement in patient comfort and motion artifacts, or to push towards lower radiotracer doses. It also opens the possibilities for dual tracer studies, more frequent follow-up examinations, and new imaging indications. How to assess quality and the potential effects of bias in training and testing sets will be discussed.

Conclusion Harnessing the power of these new technologies to extract maximal information from hybrid PET imaging will open up new vistas for both research and clinical applications with associated benefits in patient care.

Keywords PET · PET/CT · PET/MRI · Artificial intelligence · Deep learning

Introduction

Over the last 5 years, there has been an explosion of research and applications using AI methods in a wide variety of fields [1–3]. Some of the most dramatic have been in the area of medical imaging, where reports have shown that computer vision techniques built on deep learning algorithms can perform many tasks that have been thought to require human interpretive skills. These include grading of diabetic retinopathy changes on ophthalmological imaging [4], skin cancer malignancy assessment with light photography [5], bone age estimation on hand radiographs [6], and abnormality detection on frontal chest X-rays [7]. These applications are still in their

earliest stages of technical development, but there is much enthusiasm suggesting that they may lead to improved imaging diagnoses with less inter-observer variability. Some researchers in the AI field have even suggested that radiologists will no longer be needed in the near future [8].

Alongside these fascinating applications, there has been steady progress using similar AI methods to improve the quality of medical imaging [9]. Rather than focusing on replacing radiologists and nuclear medicine physicians in the interpretive space, the aim of these technologies is to create better quality images, often starting from sub-optimal or low-dose scans. Interestingly, these image-quality enhancement algorithms share many features with the computer vision applications mentioned above, relying on the power of deep convolutional neural networks. Rather than predicting diagnoses, these networks predict images; images that might be expensive to acquire or limited to non-acute patients, images that must be acquired in a shorter time interval to avoid motion artifacts, or images that need to be created at lower radiation dose than existing methods. All of these methods share the idea of "A-to-B" image transformation, where "A" is an image you have access to, and "B" is the image

✉ Greg Zaharchuk
    gregz@stanford.edu

1   Department of Radiology, Stanford University, Stanford, CA, USA

you would like to have. With enough "A–B" pairs, it is possible to train a deep neural network to learn the best possible transform from one to the other, leveraging information in the "A" image that can be used to predict the "B" image.

Hybrid imaging is well suited to provide data to train such networks. The need for simultaneous imaging is critical, since many physiological changes can occur between imaging sessions such that prediction might fail [10]. The advent of simultaneous PET/MRI in 2011 has ushered in a new paradigm for acquiring such useful information, and this article will emphasize advances based on AI that can lead to faster and safer PET images, better PET quantification, and potentially the synthesis of PET contrasts directly from MR images.

Early applications to CT and MRI

Some of the first applications of this technology for image transformation were in the area of dose reduction for CT imaging. Chen et al. demonstrated in 2017 that a residual encoder–decoder convolutional neural network (RED-CNN) could be used to denoise CT images [11]. An encoder–decoder network has two parts. The encoder is the process of taking an image and representing it with a smaller number of relevant features (somewhat like primary component analysis [PCA]) — this is commonly used for image compression. A decoder reverses this process to obtain the original image from the small number of features. The term 'residual' means that the goal is not to predict the de-noised image from the network, but rather what needs to be added to each voxel in the noisy image to create the low-noise image, which is known as the residual. The CNN is a multi-layer system of inter-connected neurons used to create the small

number of features and to restore the image quality (Fig. 1). Noisy images were created by adding synthetic Poisson noise to raw sinogram data. Training was performed by using the original images as the reference standard for prediction. This idea of "paired" image sets for the use of supervised deep learning algorithms is central to many of the algorithms to be discussed in this article. They showed that the performance of the RED-CNN was equal or superior to other state-of-the-art denoising algorithms, using the three quantitative metrics: peak signal to noise ratio (PSNR), root mean squared error (RMSE), and structural similarity index metrix (SSIM) with respect to the higher dose images. This study was an inspiration for future studies examining denoising of medical imaging. Many subsequent studies have used variations on the RED-CNN methodology and the use of patch-based training, which allows for significant augmentation of the limited amount of data available in many medical imaging settings. Due to the visual appearance of the structure of the RED-CNN, it has been described in many future studies as a "U-net" architecture [12].

Soon thereafter, Zhu et al. demonstrated the extreme flexibility of a CNN-based method for image reconstruction with a technique they called "AUTOMAP" [13]. In this paradigm, nothing is assumed in terms of the relationship between the raw sensor data (such as MRI k-space data) and the final image. If enough pairs of sensor data and associated images are available, a fully connected CNN can be used to create a method for reconstruction of subsequent images. They trained their algorithm in part on the large ImageNet dataset of images that were transformed into frequency space to simulate MRI raw data. They demonstrated that the algorithm could reconstruct MR images from various acquisition strategies (i.e., Cartesian, spiral, etc.) and showed improved performance
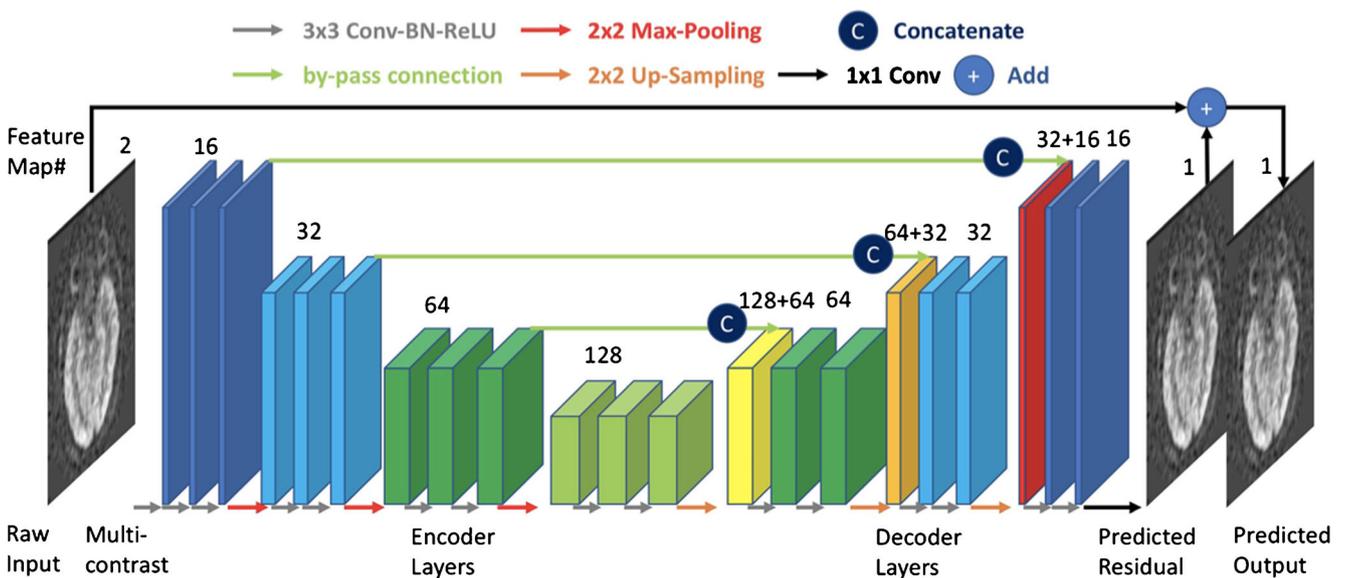

Fig. 1 Example of deep learning encoder–decoder architecture. This deep convolutional neural network (CNN) structure has been termed a "U-net" given its shape, where there is downsampling first in the image size during the encoder layers and upsampling later in the decoder layers. Skip connections between layers are used to retain spatial information

with undersampled data, with superior noise performance and immunity to artifacts, again using quantitative metrics such as RMSE. As part of this study, they also demonstrated that it was possible for a CNN to "learn" the transform to convert PET sinogram data into image space. The use of fully connected layers makes AUTOMAP challenging to apply to smaller datasets, due to the problems of overfitting and the high computational needs. More recently, Häggström et al. have shown more extensive work reconstructing PET images from sinogram data using a more computationally efficient encoder–decoder, highlighting a 100-fold speedup for reconstruction compared to standard iterative techniques such as ordered subset expectation maximization (OSEM) [14]. Other groups have shown that encoder–decoder networks and variational networks can give excellent MR image reconstruction from undersampled data [15, 16]. The importance of this work for MRI lies in its ability to markedly speed up imaging acquisition, which is useful for increasing access to scanners or for patients who cannot tolerate the length of a standard imaging examination.

MRAC studies

In the area of hybrid PET imaging, one issue that has been frequently called out is that PET/MRI is limited because of errors associated with attenuation correction, since a CT scan is not obtained [17]. One major problem is that bone and air have a roughly similar signal on most MR pulse sequences, despite having markedly different attenuation characteristics. This is of critical importance for quantitative techniques such as PET, since relative measurements are not always adequate, and even if they are, errors in a so-called "reference" region may bias measurements performed in the regions of interest. For example, in brain PET/MRI, early work demonstrated that there were systemic errors in the PET SUV values of the cerebellum, because the existing MR attenuation correction (MRAC) algorithms could not account for the large amount of bone surrounding this structure [18]. Various methods based on advances in MR imaging, such as short- or zero-echo-time imaging which can give signal from bone, have shown improvements and have been implemented by the major vendors [19–21]. Despite this work, including a landmark study of 11 different non-deep learning methods showing that most methods have biases of less than 5% for SUV [22], there is still hesitation in moving from PET/CT to PET/MRI for large-scale studies, such as clinical trials.

Several investigators have shown that it is possible to use deep learning to synthesize a "pseudoCT" from MR-only images that can be used for AC. Several groups have compared differences in SUV for PET scans corrected with a separately acquired low-dose CT or with a deep-learning based pseudoCT [23–27]. The differences from the ground truth

were negligible, such that MRAC for brain studies can be effectively considered solved. It will be important for vendors to implement these advances into their software such that studies can be compared across scanners and sites. It should be noted that the ability of using such algorithms in the rest of the body (compared with the brain) has been limited, since comparison to a "ground-truth" CT-based AC is challenging when separate imaging sessions (CT and PET/MR) are required. In particular, for abdominopelvic studies, the normal motion of bowel gas and subtle changes in patient positioning make it difficult to demonstrate the performance of the deep-learning algorithms. Some very recent work has begun to explore if the MR or CT images are required at all for predicting a pseudoCT for MRAC. Liu et al. showed recently that for brain, it is possible to train a network from 18F-fluorodeoxyglucose (FDG) non-attenuation corrected images to synthesize a pseudoCT that can be used for AC, obviating the need for a second modality [28]. Such a method could allow much simpler "PET only" devices to be constructed, potentially reducing the cost and complexity of the scanner.

Low-count PET imaging

PET offers the ability to probe molecular processes in living human beings. However, because it requires radiation, it is often perceived to be an image modality that should be reserved for older or terminally-ill patients and not used for follow-up studies, despite the tremendous value it adds due to its high molecular sensitivity. For this reason, it would be helpful to reduce the required dose to increase the number of potential use cases. Since previous work has shown the ability to de-noise another radiation-based technique (CT) [11], it is logical to examine whether it might be possible to apply this to PET as well. There are several appealing features to applying deep-learning methods to de-noise PET images: the ability to trade off dose versus bed time depending on applications, the ability to synthesize noisy images in a realistic manner using list mode data, and the ability to use hybrid imaging modalities such as CT and MR to improve the de-noising process. An example of an FDG PET scan with increasing count undersampling that can be used to create a synthesized low dose image is shown in Fig. 2.

The first work to demonstrate a CNN for PET denoising was performed by Xiang et al [29] In this study, the authors demonstrated that a low-dose PET scan obtained from 75% shortened bed duration reconstruction could be used along with a T1-weighted image as inputs to a deep learning model to predict full-dose PET images. This methodology could be used in theory to reduce scan time or dose by 4-fold. The application to dose is especially realistic, and in fact probably underestimates performance, since as dose decreases, the true
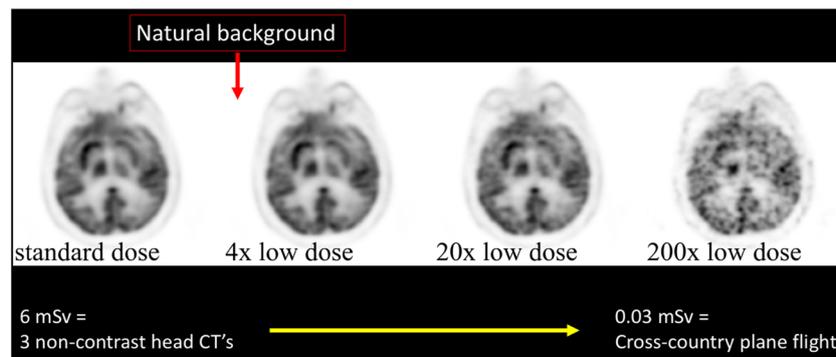
Fig. 2 Synthesized low-dose brain FDG PET imaging using undersampling from list mode data. Such methods are a conservative way to estimate images produced with lower injected doses and can be used to train a deep network to predict the full-dose PET. At 200× reduced dose, a typical FDG PET brain examination could be performed at a radiation dose comparable to the excess amount of radiation one would accrue on a cross-USA airplane flight

events in the camera scale linearly while the randoms go down in a quadratic relationship.

Xu et al. demonstrated that the inclusion of multiple MR image contrasts into the inputs for deep learning models is incredibly valuable, showing good performance for far more significant dose reductions, up to 200× for brain FDG PET/MRI studies [30]. They showed that the inclusion of MR was essential for depicting high-resolution features that could be blurred if only the low-dose PET was used as input. The performance of their network was far better than existing de-noising methods, and clinically was impactful as measured by clinical evaluations such as image quality and lesion ROI contouring. For comparison, the radiation dose of such an examination is equivalent to the excess radiation that is accrued on a single cross-USA airplane flight. Such dose reduction offers the potential to fundamentally transform how PET is used clinically.

Chen et al. applied these basic principles to amyloid brain PET imaging, which is used in the clinical assessment of dementia [31]. Amyloid normally builds up many years before the clinical symptoms of Alzheimer's disease become evident, and the absence of amyloid activity in the brain cortex can be used to rule out the diagnosis. They are typically read out clinically as positive or negative, based on whether the amyloid uptake extends into the cortex. In this study, a cohort of normal subjects and those with various dementias were scanned with PET/MRI using the tracer 18F-florbetaben. The list mode data from these scans was undersampled by removing 99 of every 100 events, leading to a simulated 1% dose image. This image, along with standard MR imaging sequences such as T1, T2, and T2 FLAIR, were used as inputs to a residual U-net CNN with skip connections, where the output image to predict was the full-dose PET image. Training was performed on both positive and negative scans. Both a PET-only and a PET/MRI model was trained, and they showed that the PET model had similar image quality metrics, such as peak SNR and structural similarity index metric (SSIM) to the PET/MR model (Fig. 3). However, the addition of the MR information was critical to resolve the high-resolution features to distinguish whether the amyloid signal extended into the cortex. Readers of the PET/MR model based on the 1% dose were about 90% accurate in their readings (positive/negative) compared to the full-dose gold standard. These full-dose images were read by the same clinicians after a washout period and had the same accuracy, showing that the performance of the CNN method was equivalent to the intra-reader accuracy of typical readers. Quantitative SUV measurements also showed marked improvement after the CNN image enhancement. More recent work suggests that combining a generative adversarial network (GAN) with feature matching into the discriminator can lead to similar performance even without the MR images, and work in this area is ongoing [32].

Image synthesis — towards "zero dose"

Lower radiation dose enabled by AI has shown great promise, but it is an open question whether PET images can be predicted based solely on data from other imaging modalities. For example, while we believe that PET yields molecular or functional information that is fundamentally not accessible to other imaging techniques, often a trained radiologist or nuclear medicine can make an educated guess about expected findings on a PET scan given the clinical history and ancillary imaging such as MRI. This suggests that in some cases, the information existing before the PET scan itself could be leveraged to improve the quality or even outright predict the images from the resultant PET study. Deep learning offers the potential to identify possible correlations that might be invisible to the clinician that could be used for such prediction.

One simple example is the predicting of cerebral blood flow (CBF) from MR imaging. Several MR techniques have been proposed and partially validated, including bolus contrast dynamic susceptibility (DSC) and arterial spin labeling (ASL). These methods often work well in healthy subjects with normal arterial arrival times but fail in unpredictable ways in older patients and patients with cerebrovascular disease. For this
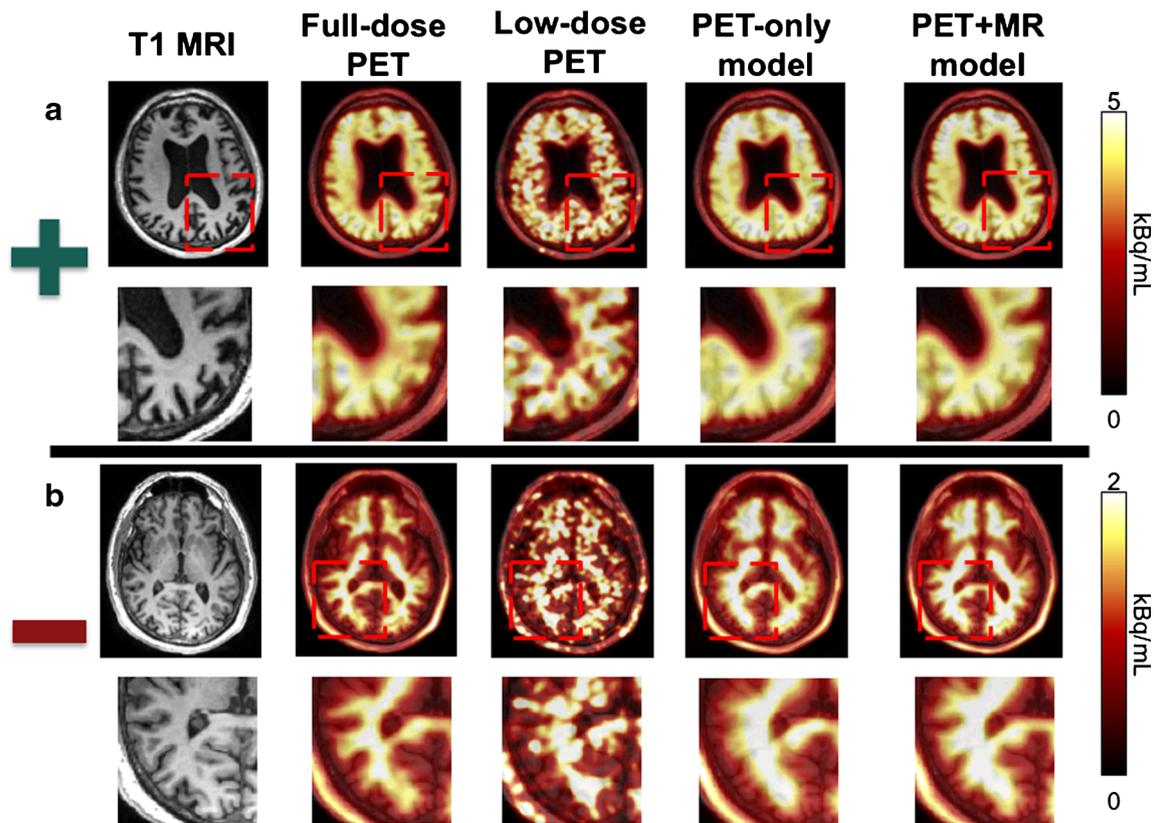
Fig. 3 Example of simulated 1% amyloid PET/MRI imaging in a patient with a **a** positive and **b** negative 18F-florbetaben amyloid scan. A deep CNN is trained on both positive and negative scans to reproduce full-dose images using either 1% dose PET only or using 1% dose PET plus several MRI sequences (T1-, T2-, and T2 FLAIR-weighted imaging). The PET-only model demonstrates denoising but at the cost of poor spatial resolution. The PET + MRI model was able to denoise but also to retain the high-resolution information that is critical to amyloid brain PET interpretation

reason, oxygen-15 water PET remains the reference standard for the CBF measurement. Using simultaneous PET/MRI, it is possible to obtain synchronous reference standard CBF measurements along with a wealth of MR functional and structural information. Given that CBF fluctuates significantly even within a single day, this is an application in which simultaneity is critical, requiring the use of hybrid imaging.

Guo et al. have used such an experimental setup to study the relationship between MR images and reference standard oxygen-15 water PET CBF maps [33, 34]. Multiple MR sequences, including structural (T2-, T2 FLAIR-, and T1-weighted imaging) and functional (many "flavors" of ASL including images of CBF and arterial delay maps) are obtained. The images are then placed into a standard template space and the network is also given access to the voxel location for learning. The image to predict at the output of the network is the reference standard PET CBF image. By training on both normal subjects and patients with Moyamoya disease, who have chronic ischemia due to severe stenosis of intracranial vessels, the group showed that it is possible to create more accurate CBF maps from MR inputs alone using the deep CNN than using any of the MRI sequences specifically designed to measure CBF alone.

While CBF is a functional measurement, the question arises whether molecular images might also be produced using such an approach. On the face of it, it seems highly unlikely that a network could be trained to approximate a specific PET tracer from MR images alone. However, it is clearly a testable hypothesis. Some initial work shows that it is possible to create pseudo-FDG images with more similarity to the true FDG image than any of the MR inputs to the model [35]. Some of this relies on the close ties between CBF and metabolism in the human brain. As with CBF prediction, it will be important to have simultaneous imaging, as pathophysiology can change quickly in many disease (Fig. 4). Another example is the work of Wei et al., who showed that an adversarial DL network based on multi-modal MRI only could accurately predict myelin content measured with ^{11}C-Pittsburg Compound B (PIB) [36].

How many datasets are required for training?

Deep learning is known to work best with very large datasets, which often number in the millions (such as ImageNet with > 14 M images). It is important to realize that tasks that require
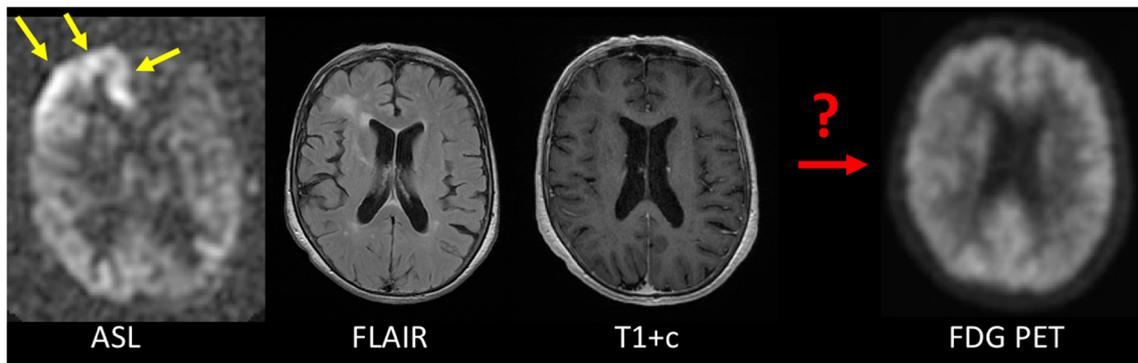
Fig. 4 An example of the importance of simultaneity for PET/MRI: a 72-year-old female with Stage IV lung cancer and altered mental status. The MRI was performed about 1 week prior to the PET scan. During the MRI scan, the patient has marked CBF increase in the right frontal lobe consistent with a peri-ictal state (*yellow arrows*). This was controlled by the time of the FDG PET scan, which shows normal metabolism in this region. This illustrates that simultaneity is important for the goals of image prediction, since the assumption that the patient is in the same physiological state cannot be made

classification of images are very different from the applications described here surrounding image quality improvement and image-to-image transformation. In the former, only one classification is available per image, such that the network can only learn relatively little from each case. In contrast, for image transformation tasks, each voxel of the image can contribute to the loss function for training. Furthermore, much imaging data is multi-slice and often independent predictions can be made from different slices, resulting in significant data augmentation. Based on the studies already in the field, it is clear that meaningful results can be obtained even with small datasets, numbering in the 10's to 100's. However, the size of the dataset needed will be very dependent on how easy the task is, so no definite rules can be given. One caveat, discussed in the next section, is that even if good results can be obtained using small numbers of subjects, there is the risk that a lack of diversity in the training data may lead to a "brittle" solution, where algorithms perform well on a "leave-one-out" type validation, but performance diminishes rapidly when exposed to larger datasets that are not well represented in the training set, sometimes known as bias.

Problems of bias in AI

Recently, more attention is being paid to potential biases in AI algorithms in non-imaging domains, such as face recognition and image search. As AI moves from being "proof of concept" in studies performed under very controlled conditions at single institutions into more widespread use, these challenges are becoming more acute. A tenet of AI prediction is that it is very dependent on the data being used to create the algorithm. It is important to ensure that this training data is representative of the diseases being assessed, or there is a risk that it will fail or underperform when exposed to new test data with different characteristics.

This was tested as part of the PET CBF prediction study described above. Networks were trained on subsets of the data consisting of either healthy control subjects or patients with cerebrovascular disease, and then tested on either representative (similar) or unrepresentative (dissimilar) cohorts. It was found that the networks trained and tested in the same group (healthy vs Moyamoya disease) performed best, while there was a dropoff in performance in the other cases. This was largest for predicting CBF in patients with cerebrovascular disease using a network trained only with normal subjects. It was found that the cerebrovascular patients had both normal and abnormal regions, and therefore the networks were better able to predict CBF in healthy subjects. Overall, the largest training set consisting of both normal subjects and patients with disease performed better than either of the individual smaller datasets, speaking to the value of larger datasets for improving performance, but that these larger datasets needed to be as representative of the population that the algorithm will be used on as possible. While this study demonstrates this effect in a small, controlled population, other more subtle forms of bias may become entrenched if AI scientists and physicians do not work together to ensure that representative populations are used for training. Early work on transfer learning, in which one network trained in a different group of patients is improved by training on a smaller set of more representative patients, is ongoing and important.

How should these new technologies be assessed?

A fundamental question is what metrics should be used to assess the quality of the images synthesized using AI. Metrics on how to assess such images have not been established, since images are often used in different ways, and the clinical impact of a synthesized image is difficult to

generalize for a specific application. For example, in an application to determine tumor recurrence, correct prediction of the entire image with the exception of the lesion itself would be useless.

Early work has focused on objective metrics such as the PSNR, RMSE, or SSIM. All of these measures require the existence of a reference ("ground-truth") image to which the created image is compared. While in general, images with higher quantitative metrics more closely approximate the desired image, many different kinds of image degradation can lead to similar levels of these metrics, some of which might be more acceptable than others. For example, errors in prediction spread evenly over the image might be less preferable than a more accurate prediction of the underlying image with an obvious artifact, since from an interpretation standpoint, radiologists and nuclear medicine physicians are familiar with "reading through" artifacts, while more subtle changes that are clinically significant could go undetected. Some cutting-edge techniques such as those using generative adversarial networks (GANs) strive to create images that can "fool" an AI-based detector when supplied with real and predicted images. This can lead to the creation of "realistic"-looking images which may not bear a relationship to the true ground-truth image, though methods that ensure a degree of data consistency with the original data can be used to control this. Such techniques can reduce the likelihood of creating "hallucinations" not present in the true images.

Clearly, it is difficult to assess if an AI-predicted image is adequate to answer the clinical question for which a study is ordered. In addition to trying to optimize quantitative metrics, there will be a large role for radiologists and nuclear medicine physicians to qualitatively assess the images within a given clinical context, to determine if similar reports and recommendations are made on the AI-enhanced images as on the ground truth images. Assessment of image quality using subjective scales (such as a Likert scale) is another way of guiding imaging scientists to produce clinically acceptable images, though this is inherently a subjective process. One significant challenge to clinical assessment of hybrid AI images is the sheer cost and complexity of performing such studies, which rely on time-poor highly-compensated specialists. However, attention to clinical features will be critical to judge the suitability and guide the development of AI-based image transformations.

Vision of the future of hybrid imaging

The amount of progress made in the last several years in the upstream area of hybrid imaging transformation has been impressive, even with only a small fraction of practicioners working in this field. It is very likely that such technologies will become ever more common moving forward. This is driven in part by the increasing collaboration between data scientists, computer scientists, and imaging professionals, each of whom bring important perspectives to the field. Computer resources, both on the hardware and the software side, are becoming more ubiquitous, with much code being released freely using online repositories. A limiting feature of the current environment is the difficulty of data sharing between institutions, since access to large, representative datasets are critical to moving the field forward and avoiding issues of bias related to smaller, single-institution studies. Given the high data density of hybrid imaging with PET/CT and PET/MRI, it is likely we will continue to see advances in this area, and there will be a strong focus on generalizability between institutions and cohorts. It is important that we as a field take our responsibility seriously, with not only an open mind but also a critical eye, with regard to the potential benefits these new technologies will bring.

Compliance with ethical standards

References

1. Krizhevsky A, Sutskever I, Hinton G. Imagenet classification with deep convolutional neural networks. In: NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1. Lake Tahoe, Nevada. December 03–06, 2012; 1097–1105.

2. Hinton G. Deep learning - a technology with the potential to transform health care. JAMA. 2018;320:1101–2.

3. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521: 436–44.

4. Grewal PS, Oloumi F, Rubin U, Tennant MTS. Deep learning in ophthalmology: a review. Can J Ophthalmol. 2018;53:309–13.

5. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542:115–8.

6. Larson DB, Chen MC, Lungren MP, Halabi SS, Stence NV, Langlotz CP. Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. Radiology. 2018;287:313–22.

7. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the chexnext algorithm to practicing radiologists. PLoS Med. 2018;15:e1002686.

8. Creative Destruction Lab. Geoff Hinton: on radiology [video file]. Retrieved from https://www.youtube.com/watch?v=2HMPRXstSvQ&t=2s. 2016. Accessed 28 Apr 2019.

9. Zaharchuk G, Gong E, Wintermark M, Rubin D, Langlotz CP. Deep learning in neuroradiology. AJNR Am J Neuroradiol. 2018;39:1776–84.

10. Fan A, Jahanian H, Holdsworth S, Zaharchuk G. Comparison of cerebral blood flow measurement with [15O]-water PET and arterial spin labeling MRI: a systematic review. J Cereb Blood Flow Metab. 2016;36:842–61.

11. Chen H, Zhang Y, Kalra M, Lin F, Chen Y, Liao P, et al. Low-dose CT with a residual encoder-decoder convolutional neural network (RED-CNN). ArXiv e-prints; 2017. p. 1702.

12. Ronnenberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: Proceedings of MICCAI, vol. 9351; 2015.

13. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature. 2018;555:487–92.

14. Haggstrom I, Schmidtlein CR, Campanella G, Fuchs TJ. DeepPET: a deep encoder-decoder network for directly solving the PET image reconstruction inverse problem. Med Image Anal. 2019;54:253–62.

15. Mardani M, Gong E, Cheng J, Vasanawala S, Zaharchuk G, Xing L, et al. Deep generative adversarial networks for compressed sensing MRI. IEEE Trans Med Imaging. 2019;38:167–79.

16. Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, et al. Learning a variational network for reconstruction of accelerated MRI data. Magn Reson Med. 2018;79:3055–71.

17. Wagenknecht G, Kaiser HJ, Mottaghy FM, Herzog H. MRI for attenuation correction in PET: methods and challenges. MAGMA. 2013;26:99–113.

18. Su Y, Rubin B, McConathy J, Laforest R, Qi J, Sharma A, et al. Impact of MR-based attenuation correction on neurologic PET studies. J Nucl Med. 2016;57:913–7.

19. Wiesinger F, Bylund M, Yang J, Kaushik S, Shanbhag D, Ahn S, et al. Zero TE-based pseudo-CT image conversion in the head and its application in PET/MR attenuation correction and MR-guided radiation therapy planning. Magn Reson Med. 2018;80:1440–51.

20. Leynes AP, Yang J, Shanbhag DD, Kaushik SS, Seo Y, Hope TA, et al. Hybrid ZTE/Dixon MR-based attenuation correction for quantitative uptake estimation of pelvic lesions in PET/MRI. Med Phys. 2017;44:902–13.

21. Ladefoged CN, Benoit D, Law I, Holm S, Kjaer A, Hojgaard L, et al. Region specific optimization of continuous linear attenuation coefficients based on UTE (RESOLUTE): application to PET/MR brain imaging. Phys Med Biol. 2015;60:8047–65.

22. Ladefoged CN, Law I, Anazodo U, St Lawrence K, Izquierdo-Garcia D, Catana C, et al. A multi-Centre evaluation of eleven clinically feasible brain PET/MRI attenuation correction techniques using a large cohort of patients. Neuroimage. 2017;147:346–59.

23. Bradshaw TJ, Zhao G, Jang H, Liu F, McMillan AB. Feasibility of deep learning-based PET/MR attenuation correction in the pelvis using only diagnostic MR images. Tomography. 2018;4:138–47.

24. Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB. Deep learning MR imaging-based attenuation correction for PET/MR imaging. Radiology. 2018;286:676–84.

25. Leynes AP, Yang J, Wiesinger F, Kaushik SS, Shanbhag DD, Seo Y, et al. Zero-echo-time and Dixon deep pseudo-CT (ZEDD CT): direct generation of pseudo-CT images for pelvic PET/MRI attenuation correction using deep convolutional neural networks with multiparametric MRI. J Nucl Med. 2018;59:852–8.

26. Gong K, Yang J, Kim K, El Fakhri G, Seo Y, Li Q. Attenuation correction for brain PET imaging using deep neural network based on Dixon and ZTE MR images. Phys Med Biol. 2018;63:125011.

27. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. Med Phys. 2017;44:1408–19.

28. Liu F, Jang H, Kijowski R, Zhao G, Bradshaw T, McMillan AB. A deep learning approach for (18)F-FDG PET attenuation correction. EJNMMI Phys. 2018;5:24.

29. Xiang L, Qiao Y, Nie D, An L, Wang Q, Shen D. Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI. Neurocomputing. 2017;267:406–16.

30. Xu J, Gong E, Pauly J, Zaharchuk G. 200x low-dose PET reconstruction using deep learning. ArXiV. 2017:1712.04119.

31. Chen KT, Gong E, de Carvalho Macruz FB, Xu J, Boumis A, Khalighi M, et al. Ultra-low-dose (18)F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs. Radiology. 2019;290:649–56.

32. Ouyang J, Chen K, Gong E, Pauly J, Zaharchuk G. Ultra-low-dose PET reconstruction using generative adversarial network with feature mapping and task-specific perceptual loss. Med Phys. 2019 (in press).

33. Guo J, Gong E, Fan A, Khalighi M, Zaharchuk G. Improving ASL CBF quantification using multi-contrast MRI and deep learning. In: Proceedings of American Society of Functional Neuroradiology; 2017.

34. Guo J, Gong E, Goubran M, Fan A, Khalighi M, Zaharchuk G. Improving perfusion image quality and quantification accuracy using multi-contrast MRI and deep convolutional neural networks. In: Proceedings of ISMRM; 2018. p. 310.

35. Gong E, Chen K, Guo J, Fan A, Pauly J, Zaharchuk G. Multi-tracer metabolic mapping from contrast-free MRI using deep learning. In: Proceedings of ISMRM workshop on machine learning; 2018.

36. Wei W, Poirion E, Bodini B, Durrleman S, Ayache N, Stankoff B, et al. Learning myelin content in multiple sclerosis from multimodal MRI through adversarial training. In: Proceedings of MICCAI; 2018. p. 514–22.