

Received:
28 March 2023

Accepted:
20 June 2023

Published online:
04 September 2023

© 2023 The Authors. Published by the British Institute of Radiology under the terms of the Creative Commons Attribution 4.0 Unported License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Cite this article as:

Reader AJ, Pan B. AI for PET image reconstruction. *Br J Radiol* (2023) 10.1259/bjr.20230292.

AI IN IMAGING AND THERAPY: INNOVATIONS, ETHICS, AND IMPACT: REVIEW ARTICLE

AI for PET image reconstruction

ANDREW J READER, PhD and BOLIN PAN, PhD

School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK

Address correspondence to: Prof Andrew J Reader
E-mail: andrew.reader@kcl.ac.uk

ABSTRACT

Image reconstruction for positron emission tomography (PET) has been developed over many decades, with advances coming from improved modelling of the data statistics and improved modelling of the imaging physics. However, high noise and limited spatial resolution have remained issues in PET imaging, and state-of-the-art PET reconstruction has started to exploit other medical imaging modalities (such as MRI) to assist in noise reduction and enhancement of PET's spatial resolution. Nonetheless, there is an ongoing drive towards not only improving image quality, but also reducing the injected radiation dose and reducing scanning times. While the arrival of new PET scanners (such as total body PET) is helping, there is always a need to improve reconstructed image quality due to the time and count limited imaging conditions. Artificial intelligence (AI) methods are now at the frontier of research for PET image reconstruction. While AI can learn the imaging physics as well as the noise in the data (when given sufficient examples), one of the most common uses of AI arises from exploiting databases of high-quality reference examples, to provide advanced noise compensation and resolution recovery. There are three main AI reconstruction approaches: (i) direct data-driven AI methods which rely on supervised learning from reference data, (ii) iterative (unrolled) methods which combine our physics and statistical models with AI learning from data, and (iii) methods which exploit AI with our known models, but crucially can offer benefits even in the absence of any example training data whatsoever. This article reviews these methods, considering opportunities and challenges of AI for PET reconstruction.

INTRODUCTION

Conventional implementation of image reconstruction for positron emission tomography (PET) involves writing computer code, a series of programming steps called an algorithm, based on an understanding of the physics and statistics of PET imaging. Using this understanding, the measured data from a PET scanner (e.g. sinogram data,¹ histoimages² or list-mode data³) are processed by the algorithm to form a reconstructed image. The term "image" refers to a representation of the spatiotemporal distribution of the PET radiotracer within the subject in the scanner.⁴ Usually this is a two-dimensional (2D) or three-dimensional (3D) array of pixels or voxels with values that are estimates of the radiotracer's concentration in each of those pixel or voxel regions of the scanner field of view (FOV).

Conventionally, imaging scientists assume they know a good physics model and a good statistical model for both the signal and the noise in the PET measured data. By solving a mathematical description of the measured data, they implement the solution by fixed computer code, to reconstruct an image based on these modelling

assumptions. Example reconstruction methods include iterative expectation maximisation (EM) algorithms⁵ for maximum Poisson likelihood reconstruction,^{6,7} which use a model of the imaging physics (the system or forward model⁸) for the signal, and a Poisson model for the noise. Such reconstruction methods are based on human intelligence, understanding and mathematical problem solving. In stark contrast, image reconstruction that is based on artificial intelligence (AI), such as methods using supervised deep learning⁹ (e.g. AUTOMAP¹⁰ or DeepPET¹¹), seeks instead to *learn* the sequence of steps of the computer code that is needed to reconstruct from measured data. This learning process, the training of an AI model, relies on many tens of thousands of examples, or samples, where each sample is composed of a PET measured dataset paired with the known or desired image output that is expected from that measured dataset. This approach learns the parameters of a deep network (effectively the programming steps of the computer code) by way of examples, rather than relying on an assumed model of the PET imaging process. Most research into AI for PET image

reconstruction takes place in the middle ground between these two extreme cases, such that the overall algorithm neither relies purely on modelling assumptions based on human intelligence nor purely on AI learning from example data pairs. A further burgeoning area of AI research for PET reconstruction uses AI to learn how to reconstruct even from just a single measured dataset in hand, without reference to any other external example data, as will be described later in this article. First though, it is useful to review the state-of-the-art conventional reconstruction methods, which assume that all models are known well via human intelligence.

Conventional reconstruction: filtered backprojection (FBP)

A well-known image reconstruction algorithm is filtered backprojection (FBP). FBP first filters the sinogram data and then backprojects the filtered data to obtain an image.¹² Early versions of FBP were implemented by a convolution of each projection view, followed by backprojection of each filtered projection view. As a side note, one of the very earliest uses of machine learning for emission tomographic reconstruction was based on learning the convolution kernel for the filtering process.¹³

Image reconstruction by FBP is a basic but robust method, being predictable due to its linearity, even though often producing noisy images. Noise in the reconstructed image is usually handled by a simple post-reconstruction smoothing, requiring just one adjustable parameter (the degree of smoothing) to yield easy to interpret images for radiologists, where the noise and resolution level of the images is usually visually self-evident. One of the reasons that FBP often produces noisy images is its lack of modelling of the varying noise levels in the data, when in fact the count-limited PET data do vary in statistical quality and can be well modelled by a Poisson distribution. Furthermore, FBP is limited by an overly simple model of the signal in the sinogram data—the data are modelled by the Radon or X-ray transform.¹⁴ Hence, before using FBP, the sinogram data need to be first precorrected for scattered events, random events, photon attenuation and detector normalisation.¹⁵ The subtractions arising from removing scatter and randoms, followed by the amplifications arising from attenuation correction, can be particularly problematic for increasing noise. As a result, the sinogram data have even greater variations in statistical quality, yet FBP still treats all the data as equally important, resulting in still noisier reconstructions. Statistical image reconstruction methods, reviewed next, avoid this inappropriate uniform weighting of the data.

Iterative reconstruction: maximum likelihood–expectation maximisation (MLEM) and OSEM

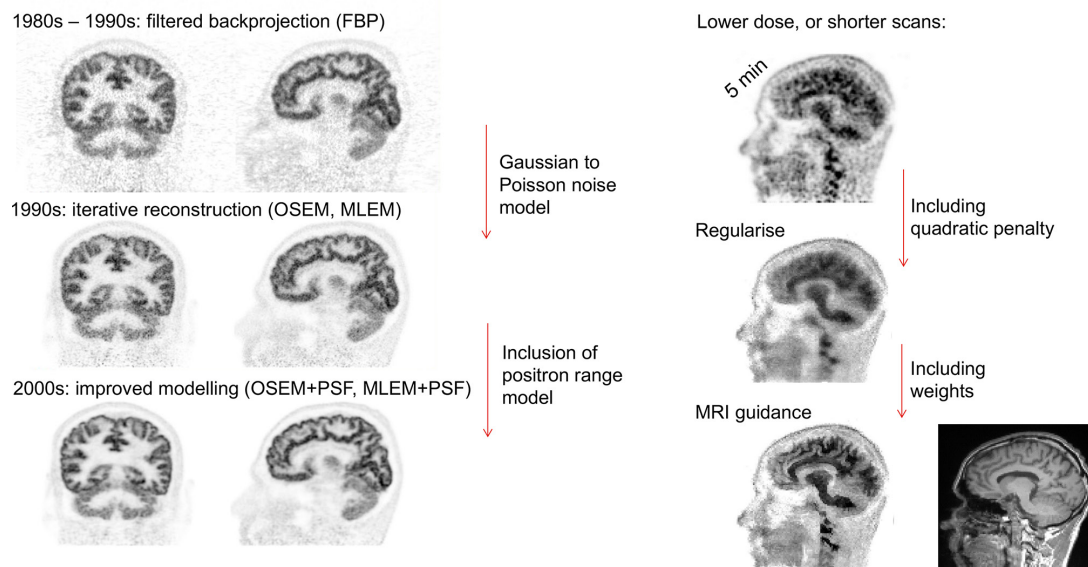
In the 1980s–90s, superior iterative methods became available for PET image reconstruction, such as the expectation maximisation (EM) algorithm for maximum likelihood (ML) reconstruction. MLEM,^{6,7} and its accelerated version, ordered subsets EM (OSEM),¹⁶ use an improved model of the noise in the data (*i.e.* Poisson counting noise, rather than Gaussian) and perhaps more importantly, allow improved modelling of the signal in

the data by using known imaging physics. Hence, rather than a simple Radon or X-ray transform model (as used by FBP, and hence requiring problematic data precorrections), MLEM and OSEM allow many physical effects, such as attenuation and normalisation, to be explicitly modelled and included within the system model of the image reconstruction algorithm.⁸ By avoiding precorrection of the data, the projection data counts remain raw and unweighted (so projection bins with low counts due to attenuation are left unamplified) and such integer (count) data are modelled well by a Poisson distribution. One example of improving a reconstruction algorithm's model of the signal in the data is the use of "PSF modelling",^{17–20} where the point spread function (PSF) is modelled during the reconstruction. By including a PSF model as part of the system matrix or forward model, spatial resolution can be improved in the reconstructed images found by MLEM+PSF or OSEM+PSF. In contrast, modelling of effects such as the PSF within FBP is not usually done (however, PSF modelling was implicit in earlier 3D backproject then filter (BPF) methods when using measured point source data).^{21,22}

Advanced regularised iterative reconstruction: maximum *a posteriori* (MAP) methods

Whilst MLEM and OSEM offer improved image quality for PET compared to FBP (as shown on the left side of [Figure 1](#)), there is nonetheless often a need to reduce scanning time and/or reduce the injected dose of the radiotracer. Consequently, noise remains an issue, and hence regularised (noise-compensated) image reconstruction methods are needed. These methods not only use: (i) an improved noise model (Poisson), and (ii) an improved model of the imaging physics, but they furthermore include (iii) a model of the type of images we expect to see, via a prior²³ or penalty. A prior assigns a probability to any given candidate reconstructed image—assigning low probabilities to images we regard as improbable for the uptake of a PET radiotracer (*e.g.* erratically noisy images), and higher probabilities to images more likely to represent a tracer's uptake (*e.g.* smoothly varying images). These regularised reconstruction methods consider both the likelihood of a candidate image given the measured data as well as the prior probability of any given image, and hence by seeking to maximise the posterior probability are called maximum *a posteriori* (MAP) methods (or, when using an EM algorithm, MAPEM methods).^{24,25} In many cases, the model of how we expect PET images to look takes the form of a penalty function rather than a probability, and such approaches should be called penalised likelihood (PL) methods. An example is Q-clear,²⁶ which assumes that the relative differences between neighbouring pixel values in an image should not be too large (as large local differences are assumed to be associated with noise rather than with clinically useful image features).²⁷ More recently, use of complementary structural information, such as from MRI, has been used to provide guidance to the regularisation,^{28–31} allowing preservation of large local differences in images that are due to legitimate image edges and features, rather than due to noise. [Figure 1](#) summarises the progress in image reconstruction from PET, starting from the 1980s, following all the

Figure 1. The progress of PET image reconstruction over the decades. On the left hand side, unregularised reconstruction methods started with 3D FBP, progressed into Poisson noise modelling with MLEM and OSEM, and then began to improve the model of the data with PSF modelling. On the right hand side (*shown for a different subject, tracer and scanner*), regularisation methods were adopted to compensate for noise, including anatomical guidance of the regularisation (e.g. via MRI). 3D, three-dimensional; FBP, filtered backprojection; MLEM, maximum likelihood–expectation maximisation; OSEM, ordered subsets expectation maximisation; PSF, point spread function.



way through to methods developed over the >40 years which followed.

Regularisation by change of spatial basis: kernel methods

One further important category of conventional methods to mention is the kernel method (KEM), proposed by Wang and Qi in 2015.³² This method amounts to a change of spatial basis functions: instead of regarding an image as a collection of independent pixels or voxels with coefficients (amplitudes or values) being found for each one, the kernel method groups pixels according to their similarity in terms of some additional features. The feature data could simply be a 3-point time–activity curve (TAC) for each pixel, or could be the registered MR image value at each pixel, or even the patch of MR pixel values that are the neighbours of the pixel.³³ By whatever means, each PET pixel has a list of additional values (a feature vector) associated with it. PET pixels can thus be grouped together into clusters, according to the similarity of these feature vectors. Conventional MLEM can then be used to find the coefficients for these pixel clusters (spatial basis functions), and the implicit regularisation achieved can be impressive: see Figure 2, where MR images were used to form the kernel spatial basis functions.³⁴

AI METHODS: FROM DIRECT TO ITERATIVE

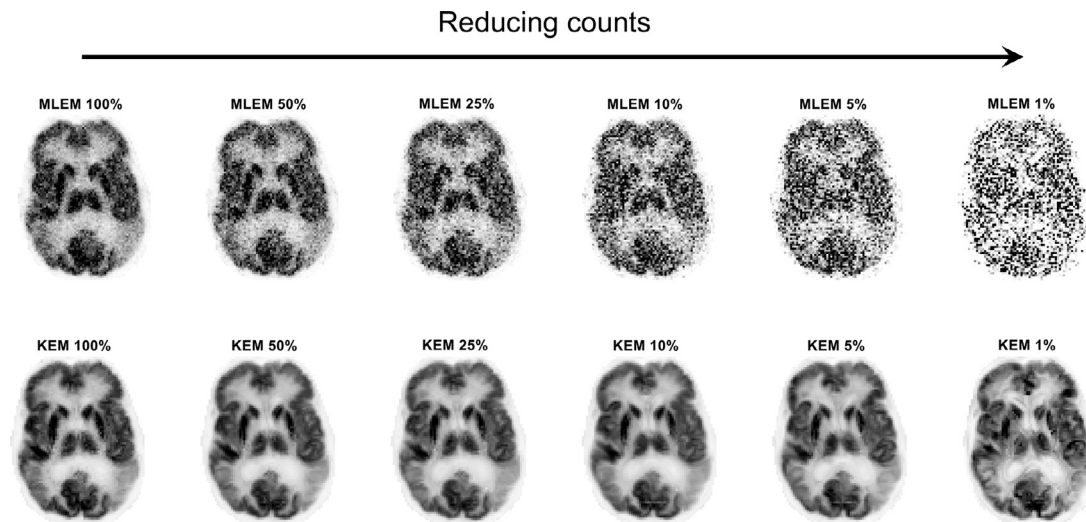
Given the sophistication of conventional state-of-the-art image reconstruction methods, one might wonder whether further improvement is possible. However, conventional approaches suffer from simplified models and assumptions. While the models of the PET scanner (captured in the forward model or system matrix) may be of acceptable accuracy, and while the Poisson

noise model may be appropriate, when it comes to modelling our expectations of the images, through a prior or a penalty, this is where conventional modelling is either too basic or too imposing. A simple noise compensation approach is to use a quadratic penalty,^{35,36} which amounts to neighbourhood smoothing of the reconstructed image values. This can be regarded as robust, but noise is reduced at the cost of spatial resolution, with edges in the images being compromised, or even small features or lesions being smoothed to become low contrast, even indiscernible, regions. If anatomical guidance information is used (e.g. from CT or MR images), this can be too imposing, as the user needs to decide how strongly to pay attention to the anatomical information. While anatomy can usefully guide the smoothing process of regularisation, it can be misleading, as there will invariably be structures and features in an MR or CT that are discrepant with those of the PET image for a given radiotracer.³⁰ As a result, even spatial resolution and edge preserving guided regularisation can give poor results in some image regions,^{36,37} through smoothing away of some PET features while also wrongly enhancing edges or structures that are not truly present in the PET data. Hence, there is clearly scope for improving conventional reconstruction, whether by improving the physics or noise modelling, or, perhaps more significantly, by improving the modelling of the penalty or the prior (the probabilistic model of how we expect PET images to appear).

Direct AI methods for PET reconstruction

To start the review of AI for PET reconstruction,³⁸ we will first consider the extreme case of ignoring nearly all modelling assumptions. Such approaches learn the entirety of the image reconstruction algorithm (*i.e.* the imaging physics, the data statistics and the prior model for the reconstructed images).

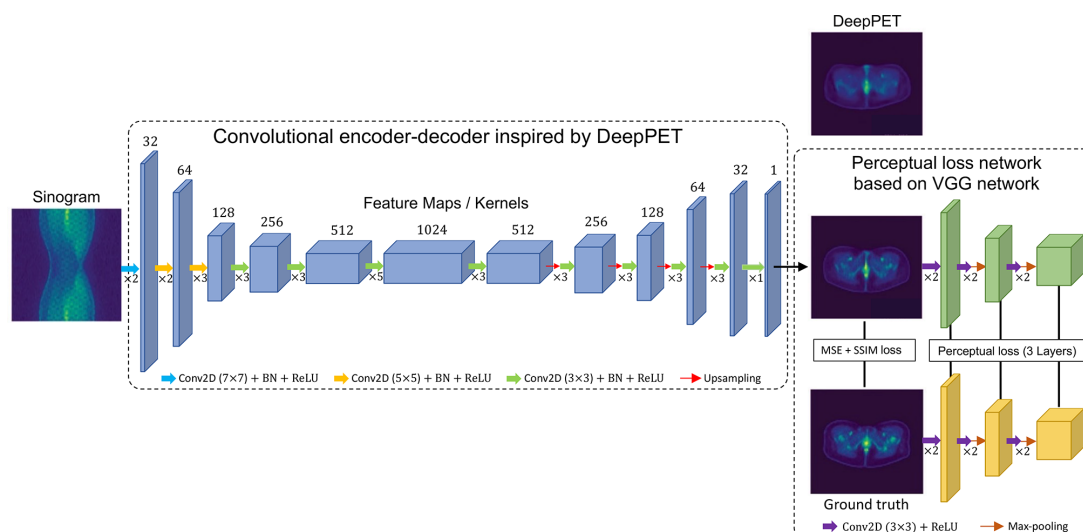
Figure 2. The kernel method for PET image reconstruction when using an MR image to provide the feature vectors for forming the spatial basis functions. This figure shows how this change of spatial basis can result in notable noise reduction even when just a small percentage of the original data are used.³⁴ The original KEM³² was based on similarities of 3-point time-activity curves for each voxel when deriving the spatial basis functions. KEM, kernel method; MLEM, maximum likelihood-expectation maximisation; PET, positron emission tomography.



Methods which seek to learn the whole mapping from measured data to a reconstructed image without use of known models will be referred to here as direct AI methods. A well-known example is DeepPET,¹¹ shown in Figure 3, with example images shown along with a more recent adaptation⁴⁰ of the approach for long-axial FOV PET imaging. The deep network starts with convolutional encoding layers, which learn to analyse the sinogram for small scale features at a given level of spatial resolution. The resulting feature maps are in turn analysed for features by further convolutional layers, to learn multiple higher-level assemblies of features, before then spatial downsampling to assist in broadening the spatial coverage of the feature learning

while still capturing the entire information from the whole input sinogram. This learned analysis of feature hierarchies is analogous to perceiving real-world objects as first composed of atoms, then of higher-level groupings of atoms (molecules), then numerous various assemblies of molecules to form materials, all the way through to larger assemblies of materials that form larger scale objects. The convolutional encoding and downsampling continues up to a latent space composed of many feature maps of very limited spatial sampling, at which stage further processing layers are learned for that low level of spatial sampling. After latent space processing, a convolutional decoder is learned for image synthesis, with progressive increases in spatial sampling

Figure 3. Representation of the performance of direct AI (purely data-driven) methods through the example of DeepPET¹¹ and recent modifications.^{39,40} The example sinogram input and DeepPET reconstructed output are courtesy of Kuangyu Shi, related to their work on improving the performance for long-axial FOV PET, with their result shown as the output of the network here (image in row 2 of the figure).⁴⁰ FOV, field of view; PET, positron emission tomography.



and corresponding decreases in the number of feature maps, to arrive ultimately at a single output high spatial resolution feature map, which is trained to be the final reconstructed image. This kind of architecture is referred to as a convolutional encoder decoder (CED).^{41,42}

By analogy to conventional Fourier-based reconstruction methods, we can broadly regard a CED as learning from data how to perform an encoding transform to a latent space (by simple analogy, Fourier transforming to the frequency domain) and then decoding from the latent space to synthesise an image (by analogy, inverse Fourier transforming the processed frequency domain data).

The supervised training of a CED relies on an optimiser (such as the gradient-based Adam algorithm)⁴³ to adapt the parameters of the encoder and decoder networks such that the output from a given input sinogram matches the associated target image for that sinogram, according to a mean square error (MSE) loss function. The target image used in supervised learning can be the ground-truth image (if known, *e.g.* by imaging known test objects, or by advanced Monte Carlo simulations)⁴⁴ or a high quality image obtained by reconstructing high count data by other methods.

Strictly this CED network is not without a model—there are constraints on the range of possible output images that can be generated from this architecture, so the network still has what is known as an inductive prior or inductive bias. As can be seen in [Figure 3](#), the output results from DeepPET can appear overly smooth, lacking in detail. Methods which followed, such as a modified implementation for long-axial FOV PET (also shown in [Figure 3](#)), and the earlier DPIR-Net,³⁹ sought to improve DeepPET by modifying the loss function used during network training. Instead of just focusing on agreement between the network output and the label or target in a training set by MSE, these later methods also required agreement at a feature map level (perceptual loss, using a pretrained VGG⁴⁵ network) as well as, for DPIR-Net, requiring the reconstruction outputs to be able to pass a discriminator test. The discriminator⁴⁶ is a network which is trained to identify real high-quality images from ones generated by a network. This discriminator can hence be used to enhance the training of the generator (if a network's output is classified as generated rather than real, this causes the optimiser to search for network parameters which generate outputs that look more real—to “fool” the discriminator). This is the well-known “generative adversarial network” (GAN) approach.⁴⁶ Hence, the trained generator network can produce reconstructed images that look indiscernibly different to real high quality images.

The long-axial FOV modified implementation reports fast reconstruction times (~20× faster), relies on single slice rebinning (SSRB)⁴⁷ to preprocess the fully 3D data into manageable 2D sinograms, and very conveniently enables scatter and normalisation corrections to be handled entirely by the learned CED network. It is however acknowledged, as visible in [Figure 3](#), that the reconstruction accuracy of DeepPET and the modifications are not quite yet at the level to meet clinical requirements, but

there is no reason why this should not be attained in future with improved training and dataset availability.

[Table 1](#) summarises a number of the current direct AI methods. The notable aspects of these direct AI methods are that: (i) the majority only operate on 2D sinograms rather than fully 3D sets of sinogram data, (ii) many tens of thousands of training pairs are needed, (iii) training can be slow, (iv) reconstruction speed is typically fast. Perhaps one of the more notable methods is that of FastPET,⁴⁸ as it makes direct use of image-space histograms. Histograms have the advantage of compressing time of flight (TOF) PET data into a 3D image-space format. By effectively compressing large TOF datasets into 3D images, relatively straightforward image-to-image deep learning mappings (such as U-Nets)⁴⁹ can be used for fast 3D image reconstruction. This method has been successfully extended and further improved,^{50,51} demonstrating good potential.

Iterative AI methods for PET reconstruction

Just as conventional reconstruction algorithms can be broadly categorised as either direct (a non-iterative sequence of programming operations) or iterative (a sequence of coding steps repeated a variable or indefinite number of times), so also there are comparable categories in deep learning. This section now covers the iterative, or unrolled, AI methods.

[Figure 4](#) shows how a conventional iterative MAPEM algorithm³⁵ can be unrolled into a sequence of modules or blocks (from left to right in the figure, one for each update of the image). Each module performs refinement of the current image estimate to improve its agreement with the data (*e.g.* based on the EM update derived for Poisson distributed data) whilst also balancing this with agreement with the prior via an analytically derived denoiser (*e.g.* derived from a quadratic penalty, which encourages images to have small local differences between pixel values). Hence, these conventional iterative methods are simply a sequence of steps, layer after layer of processing, and can be regarded as fixed, non-trainable, deep networks. The second row of [Figure 4](#) shows how the denoising (regularisation) step in MAPEM can be replaced with a trainable deep network, to learn from paired data examples how to relate noisy image updates to their high-quality counterparts. The third row of the figure shows how reference high quality reconstructions (*e.g.* from higher count sinogram data) can be used to provide training targets for the unrolled network. [Figure 5](#) shows the kind of results obtained from training an unrolled iterative method (MAPEM) with a learned regulariser, based on 45 3D training examples.

A range of unrolled reconstruction methods have now been developed, and some key methodologies are shown in [Figure 6](#). Each one performs regularisation via a learned network in different positions in the processing pipeline. Examples include gradient descent with a learned regulariser (comparable to EM-Net⁵⁴ for PET, or similar to a variational neural network (VNN) for MRI),⁵⁵ learning the gradient of the regulariser (*e.g.* FBSEM-Net),⁵² learning the proximal operator and learning to denoise (Iterative Neural Network).⁵⁶ A further example would be an unrolled alternating direction method of

Table 1. Example direct AI reconstruction methods

Name	Architecture [total parameters]	Input	Target	Loss function and optimiser (epochs) validation	Number of samples (training, validation, test)
AUTOMAP Zhu et al. 2018 ¹⁰	Two fully connected layers and CNN [~800M]	2D noisy sinograms	T1w brain images (128 × 128)	MSE with L1-norm penalty on network weights in final hidden layer RMSProp (100) Validation not used	50,000 n/a 1
DeepPET Hägström et al. 2019 ¹¹	CED [>60M]	2D noisy sinograms (269 × 288)	Ground-truth PET images (128 × 128)	MSE Adam (150) Validation used	203,305 (70%) 43,499 (15%) 44,256 (15%)
DPIR-Net Hu et al. 2020 ³⁹	CED [>60M] Discriminator [>3.5M]	2D noisy sinograms (269 × 288)	Ground-truth PET images (128 × 128)	Wasserstein GAN + VGG + MSE Adam (100) Validation not used	37,872 (80%) n/a 9468 (20%)
CED extended to SSRB sinograms from large FOV PET Ma et al. 2022 ⁴⁰	CED [~64M]	2D noisy sinogram (269 × 288)	Reconstructed PET images using OSEM + PSF/TOF from list-mode data (128 × 128)	MSE + SSIM + VGG Adam (300) Validation used	35,940 (76%) 5590 (12%) 5590 (12%)
FastPET Whiteley et al. 2021 ⁴⁸	U-Net [~20M]	Noisy histo-image slices + attenuation map slices (2 × 440 × 440 × 96)	Image slices reconstructed using OSEM + PSF (440 × 440 × 96)	MAE + MS-SSIM Adam (500) Validation used	20,297 slices (74%) 1767 slices (6%) 5208 slices (20%)

CNN refers to a convolutional neural network, CED refers to a convolutional encoder-decoder. VGG in this table refers to perceptual loss based on a VGG network.

3D, three-dimensional; FBP, filtered backprojection; MLEM, maximum likelihood-expectation maximisation; OSEM, ordered subsets expectation maximisation; PSF, point spread function.

multipliers approach, with a deep network embedded (such as deep ADMM-Net).⁵⁷

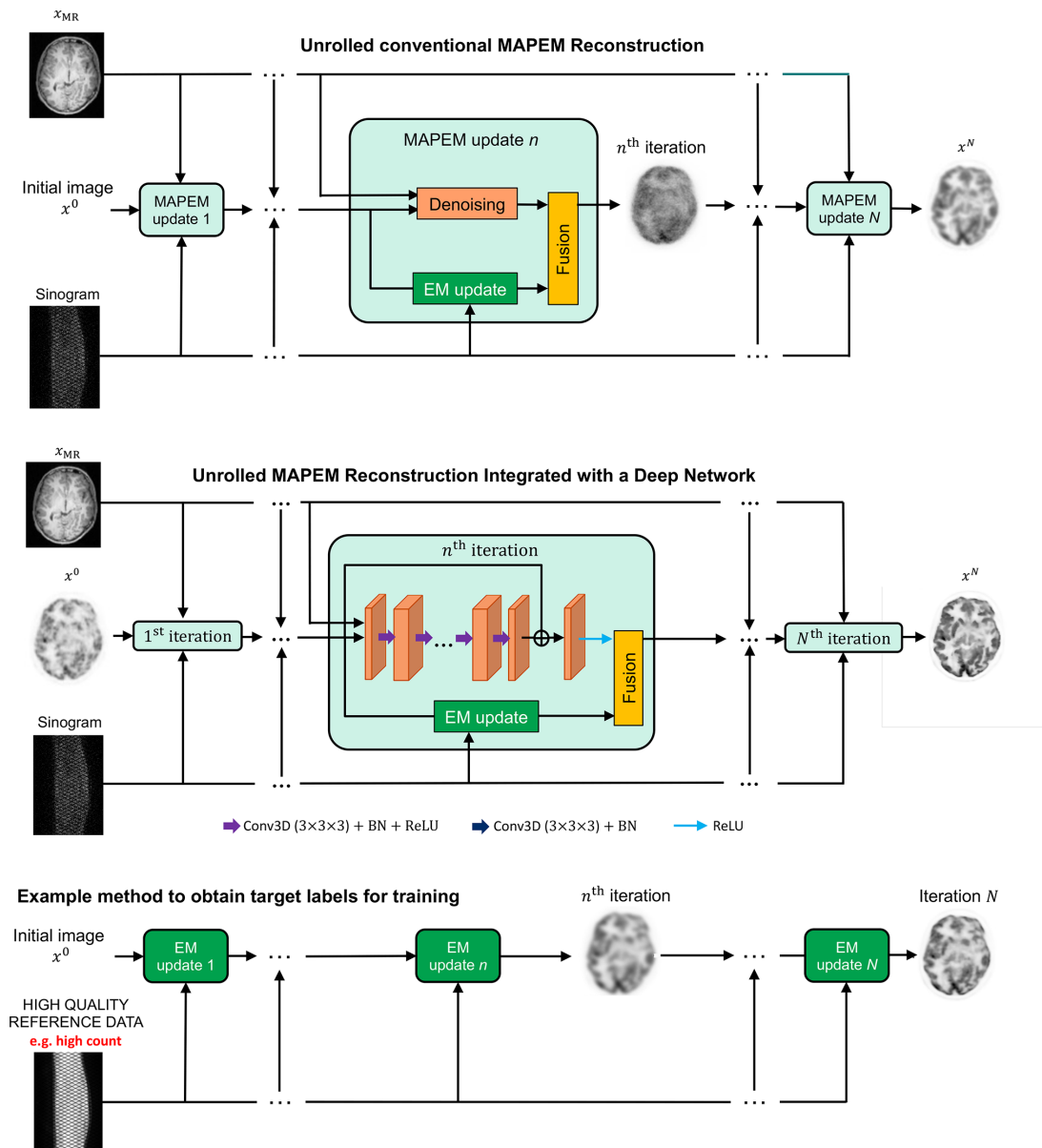
One final example of unrolled methods is the learned primal dual (LPD) method,^{58,59} based on the primal dual hybrid gradient (PDHG) method,⁶⁰ but with learning of the processing in both the primal (image) and dual (sinogram) domains, as show in Figure 7. Table 2 summarises a variety of unrolled, iterative, deep learning reconstruction methods for PET, where in one case a transformer architecture is used. Whereas CNNs often use small kernels in each network layer (hence needing many layers to encode an input to a latent representation), transformers use long-range comparisons of features across the entirety of an image in a single layer, via their attention mechanism. Transformers are broadly analogous to a trainable version of state-of-the-art conventional image processing via non-local means, where an image patch is compared with all other patches in an entire image for the purpose of contextualisation / denoising. Table 3 conveys the advantages of unrolled methods in general compared to direct AI methods: use of the physics model, less training data needed, and practical for fully 3D image reconstruction. Nonetheless, it is important to note that not only are these unrolled methods computationally intensive to train, even at run time (test time on new data) they can potentially be slow to run, being dependent on multiple iterations of forward and backprojections. It is though worth noting that differing training methods are available, some of which are more practical and less memory intensive.^{63,64}

We finish this section by mentioning some of the more advanced unrolled iterative methods, which seek to reconstruct PET and MR data simultaneously, as shown in Figure 8. This joint PET-MR unrolled iterative reconstruction method⁶⁵ learns a joint PET-MR regularisation, seeking to benefit from the common information present in both the PET and MR data, whilst preserving the crucial differences between PET and MR images.

AI METHODS WITHOUT TRAINING DATA

We have so far covered direct AI methods and iterative AI methods (which integrate deep learning into otherwise standard iterative image reconstruction schemes). Both categories rely on training data for supervised learning of the parameterised deep networks used in each type of method. A third major category of research into AI for PET image reconstruction is the case where no training data are needed at all, but instead, deep learning methodology is used either to learn a representation of the image (for just the measured dataset in hand), or to learn a dedicated reconstruction operator for the data in hand to reconstruct. Hence, no external training data are needed, and AI is exploited instead to solve processing problems. As there are no example or external training data, such methods do still need a model of the imaging physics and specification of an image reconstruction objective (e.g. ML, or MAP).

Figure 4. The example of FBSEM-Net. Row 1 shows an unrolled iterative MAPEM reconstruction, row 2 shows integration of a deep network (FBSEM-Net) and row 3 shows an example way of generating reference target labels for the network. Here, the EM update improves agreement of the image with the measured sinogram data, and can be called a data-fidelity based update. In an MRI reconstruction context, a data-fidelity update is usually achieved by a least-squares gradient descent step. EM, expectation maximisation; MAPEM, maximum a posteriori expectation maximisation.



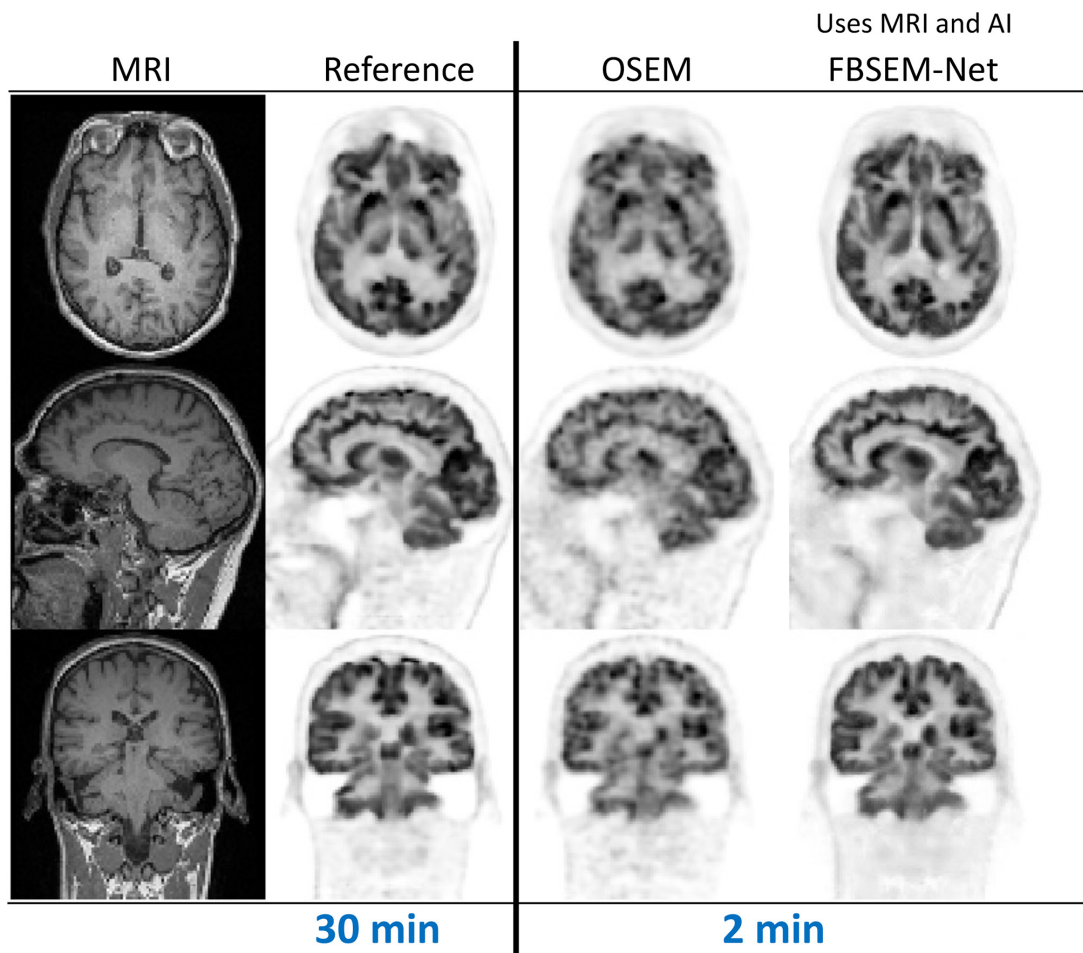
AI for image representations

The deep image prior⁶⁶ (DIP) was proposed in 2017–18, and delivered surprising results for image denoising, inpainting and super resolution in the absence of training data. The method exploits the inductive prior of a CNN, whereby the sequence of convolutional layers with non-linearities can generate, with appropriate initialisation, a range of image outputs that are regularised in the sense that they readily locate within broadly feasible manifolds occupied by medical or natural images. Typically, the noisy data are used as the target (or, for PET reconstruction, the noisy sinogram is used as the target after first applying the forward model to the generated output reconstruction). The regularisation

benefit depends entirely on the level of training—if the networks are trained towards convergence, then there will be near perfect matching of the noisy data. This still solves the inverse problem (*e.g.* for PET image reconstruction),⁶⁷ but would eliminate any regularisation benefit afforded by a CNN representation of the image.

Figure 9 shows a simple example of using the DIP for PET image reconstruction.⁶⁸ Either random noise or the patient's MR image are fed into a randomly initialised CNN, such as a U-Net.⁴⁹ The output image is then forward modelled, and the training loss function for the CNN is the MSE between the forward modelled data and the actual measured data. Hence, with increasing

Figure 5. Example results from FBSEM-Net.^{52,53} FBSEM-Net offers improved image quality when reconstructing data equivalent to just 2 min of scanning time, with image quality competitive to the reference reconstruction from ~15 times more data (30 min of data). The results shown are for “test time” reconstructions (*i.e.* for reconstruction of data not included in the training of FBSEM-Net). OSEM, ordered subsets expectation maximisation; FBSEM, forward backward splitting expectation maximisation.



optimisation of the CNN’s parameters, the network will produce an image which, when forward projected, agrees (to within the range of the forward operator) with the measured data. Early termination is therefore needed to see any regularised result. If run to convergence, the CNN representation is no different to a pixel or voxel grid representation, and the reconstruction algorithm amounts to using a deep learning optimiser (such as Adam)⁴³ to find a least-squares (or for Poisson data, a maximum Poisson likelihood) estimate. Prior even to the work of Hashimoto et al,⁶⁸ Gong had proposed a more elaborate optimisation framework using a CNN representation of the image.⁶⁹ This method was designed to avoid the need for embedding the system model into the deep network and hence avoid backpropagation of loss function gradients through the system model during training.

In seeking to resolve the issue of needing to stop the DIP training early, Bayesian methodology has been investigated by many. Examples include using a prior for the network parameters (and even using posterior distributions to quantify and make use of the uncertainty^{70,71}) and use of stochastic gradient Langevin dynamics (SGLD) for fully Bayesian posterior inference via the DIP.^{72,73} A further

approach has been Stein’s unbiased risk estimator (SURE) as the loss function for training the DIP (DIP-SURE) instead of using an MSE loss, again to avoid the overfitting problem.⁷⁴

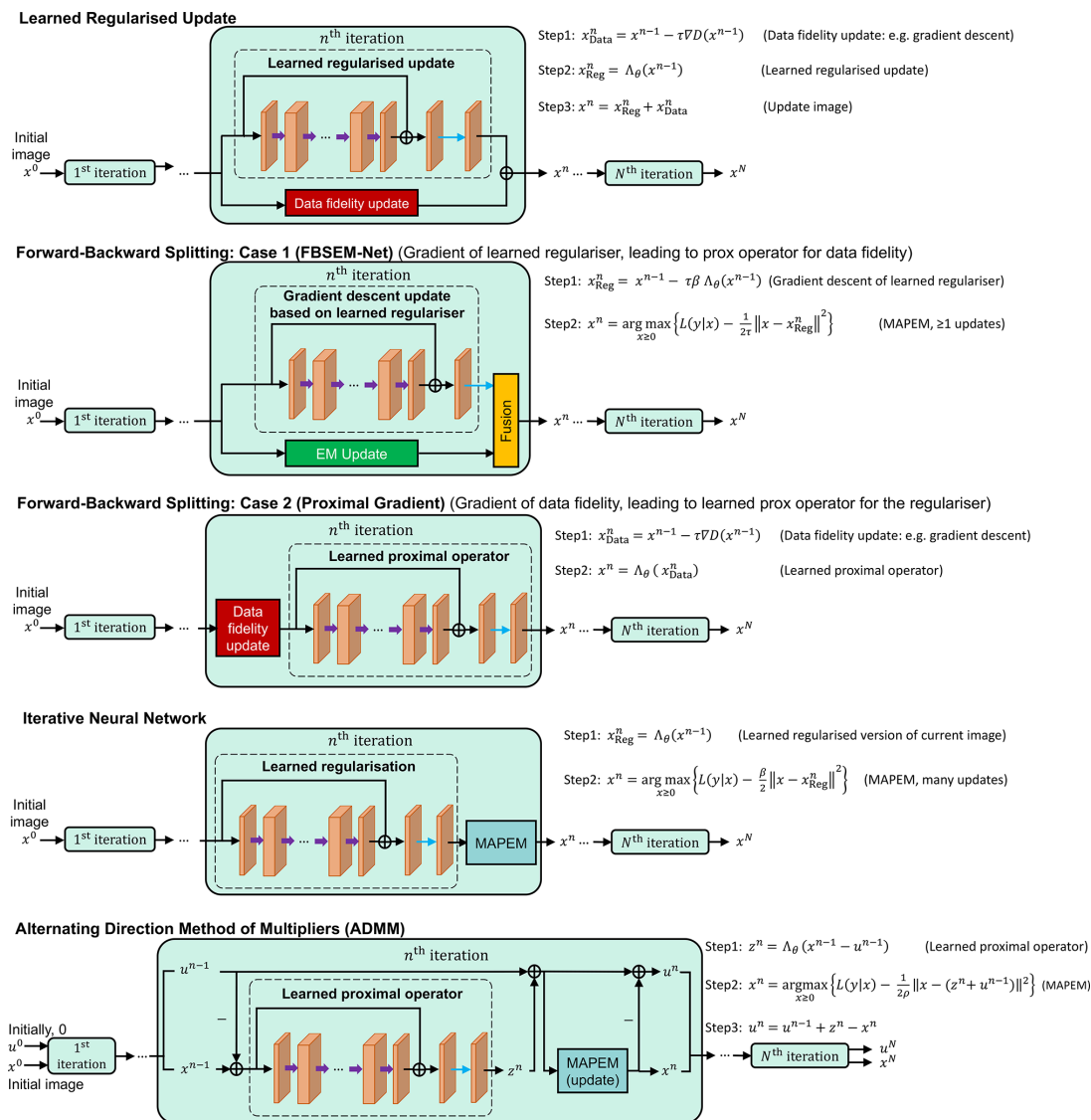
More recently, Gong et al have taken the DIP method further, by combining it with the kernel method, and then also with direct kinetic parametric estimation.⁷⁵ Figure 10 shows the impressive results achieved for direct estimation of Patlak slope (K_i) images.

Table 4 summarises the methods that exploit the DIP for PET image reconstruction, including also some AI-assisted variations of KEM, which will be covered later.

AI for reconstruction operators

An alternative use of AI in the absence of training data is to define a deep network to perform processing at certain stages in the image reconstruction pipeline. Figure 11 gives the example of deep-learned FBP (DL-FBP): a deep network is used to process a sinogram in such a way that, when the processed sinogram is backprojected, a reconstructed image is obtained. The image is a reconstruction, as the training of the deep network is based on

Figure 6. Five example unrolled iterative image reconstruction methods with deep networks embedded. Row 1 corresponds to taking a current image and updating it according to the data (a gradient descent step to increase agreement with the measured data), and also updating it according to the prior knowledge of how the image should appear (simply using a deep network to provide this increased agreement). Row 2 is very similar, but uses an analytically derived way for combining the two updates (rather than a simple addition, as in row 1). Row 3 corresponds to a deep learned version of a proximal gradient update. The example in row 4 is the sequential method but now using an analytical form of the update. Finally, row 5 shows how the ADMM method can have a network embedded, and it amounts to the same approach as in row 4, but with use of a residual image, u .



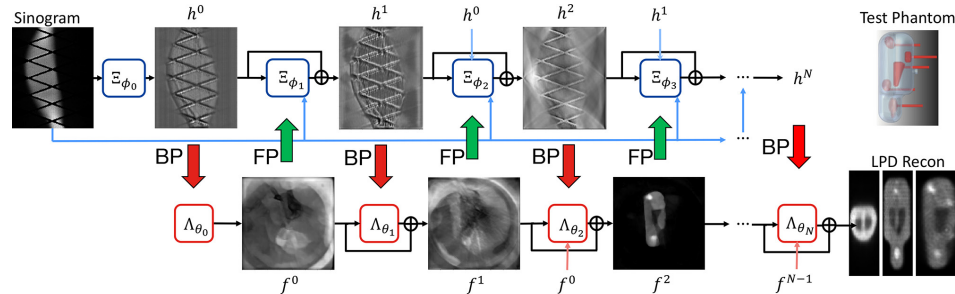
minimising the discrepancy (loss) between the forward projection of the image and the actual input measured data. This is an example of filling in a potential gap in our knowledge by deep learning; we know what we require (a reconstructed image that agrees with the sinogram data according to some reconstruction objective function), but we may not know how to do the processing of the data to achieve this. AI can be used to solve such tasks, tapping into the optimisation methods that have been so successful in deep learning. An important further development on this theme is to combine an image-space deep network with a data-space network,⁷⁹ to obtain DL-FBP-F. This approach,

just as DL-FBP, can be implemented just by self-supervised learning from the dataset in hand, or optionally combined with supervised learning. If used exclusively with supervised learning, then the DL-FBP-F method becomes similar to just one module of an unrolled method, but with much reduced computation compared to the many modules usually employed in unrolled methods, which may reduce their level of adoption.

AI-assisted KEM

A final important category, related to the example of using the kernel method with the DIP as previously shown in Figure 10

Figure 7. LPD⁵⁸ when applied to example PET phantom data.⁵⁹ Figure adapted from data courtesy Massimiliano Colarieti-Tosti. The overall processing pipeline effectively has two rails: the top rail conducts sequential processing of the sinogram data and the lower rail conducts sequential processing of the image. These two processing subpipelines are interconnected by forward and backprojection operators (FP and BP) to map between the image domain and data domain. BP, backprojection; FP, forward projection; LPD, learned primal dual; PET, positron emission tomography.



(and Table 4), is using a deep network to learn the best features to use for forming the kernel matrix used in KEM, called the deep kernel representation.⁷⁸ Training data are derived from the single dataset to reconstruct (by generating low-count versions of the data), and the method elegantly learns what are the most appropriate features to pick out from an input prior image such that when these features are used in a voxel-level similarity comparison to form spatial basis functions for KEM, the output image best matches the full-count high quality reference. These basis functions are then used with standard KEM for reconstructing the whole dataset. Example results from this deep kernel representation are shown in Figure 12, where it outperforms the conventional DIP (with early termination and oversmoothness) as well as conventional KEM based on 3-point TAC feature vectors.

OPPORTUNITIES, CHALLENGES AND SUMMARY

These significant advances in using AI for PET image reconstruction are not without concerns. A well-known concern, e.g. demonstrated for AI-based MR reconstruction, is the possibility of hallucinations (false positives) and the possibility of deletions (false negatives) from images.⁸⁰ Whereas these effects were always possible with conventional reconstruction (a noise spike looking like a lesion, or noise and smoothing seeming to remove a lesion), in conventional reconstruction it was always easy for an observer to assess the quality of an image, and hence make allowances for such occurrences. With the improved image quality from AI methods, hallucinations or deletions may not be at all appreciable based on the overall impression of image quality. A potential solution would be

Table 2. Unrolled methods for AI in PET image reconstruction.

Name	Architecture [total parameters]	Input	Target	Loss function and optimiser (epochs) validation training method	Number of samples (training, validation, test)
EM-Net Gong et al. 2019 ⁵⁴	10 modules U-Net shared [~2M]	Previous output/ iteration	3D high count reconstruction (128 × 128 × 46)	MSE Adam Validation not used Gradient truncation	16 n/a 1
MAPEM-Net Gong et al. 2019 ⁶¹	8 modules U-Net not shared [~16M]	Current output from a block	3D high count reconstruction (128 × 128 × 105)	MSE [Details not specified] Validation not used End-to-end	18 n/a 1
FBSEM-Net Mehranian and Reader 2020 ⁵²	10 modules CNN shared [~77k]	Previous output / iteration	3D high count reconstruction— cropped (114 × 114 × 128)	MSE Adam (200) Validation used Gradient truncation	45 5 5
Iterative Neural Network Lim et al. 2020 ⁵⁶	10 modules CNN not-shared [~40k]	Current output from a block	3D true activity image (200 × 200 × 112)	MSE Adam (500) Validation not used Sequential training	4 n/a 1
TransEM Hu and Liu 2022 ⁶²	10 modules Swin Transformer [details not specified]	Previous output / iteration	2D high count reconstruction	MSE Adam Validation used Gradient truncation	510 30 60

AI, artificial intelligence; CNN, convolutional neural network; 2D, two-dimensional; 3D, three-dimensional; PET, positron emission tomography.

Table 3. Broad level comparison of core approaches to using AI in PET image reconstruction.

	Direct AI methods	Unrolled AI methods	AI methods without training data
Interpretability	Low	Good	Good
Use of physics model	No	Yes	Yes
Generalisation (outside training distribution)	Can be challenging	Better than direct methods	Good
Suitability for fully 3D image reconstruction	Challenging	Yes	Yes
Typical number of training pairs	>10,000 (2D reconstruction)	~50 (3D reconstruction)	1
Inference time/time to reconstruct from new data	Fast	Potentially the slowest	Faster than unrolled methods, depending on implementation
Expected image quality if within training distribution	Potentially the best	Good	Does not benefit from population data

2D, two-dimensional; 3D, three-dimensional; AI, artificial intelligence; PET, positron emission tomography.

to take the Bayesian approaches previously described (developed also in the MR context)⁸¹ and generate additional images reflecting the level of uncertainty in the reconstructed image values. Yet that would increase the image reading burden for radiologists. Nonetheless, Bayesian deep learning approaches, or methods based on deep ensembles⁸² or even drop-out,⁸³ are ways to reveal the epistemic uncertainty in reconstructed images. Related to this is a very recent advance in generative AI modelling: diffusion models (popularised by image generation models such as Stable Diffusion).⁸⁴ These models can be trained to provide prior probabilities for high quality images,

and as such are now finding their way into image reconstruction^{85,86} and producing state-of-the-art results (as so far demonstrated for MRI reconstruction), providing also uncertainty estimates.⁸¹ Early work in PET has already begun in a post-processing context.⁸⁷

For PET image reconstruction, there also remains the need for widely accepted benchmark datasets and tests (as increasingly used in MRI),^{88,89} needed not just for conventional methods but also AI approaches. Early developments in high

Figure 8. The example of unrolled joint PET-MR reconstruction from Corda-D'Incan et al.⁶⁵ Example results are shown, comparing OSEM with MAPEM (Q, using a quadratic penalty, BQ using a Bowsher MRI-guided quadratic penalty), and with independent deep learned unrolled reconstruction from PET data only (P) and with MR guidance (P MR). The final result shown is the joint reconstruction of the PET data—revealing preservation of a PET-unique lesion (absent in the MR) whilst retaining good image quality globally—effectively seeking the best aspects of “Indep P” and “Indep P MR” simultaneously. OSEM, ordered subsets expectation maximisation; MAPEM, maximum a posteriori expectation maximisation; MSE, mean square error; PET, positron emission tomography.

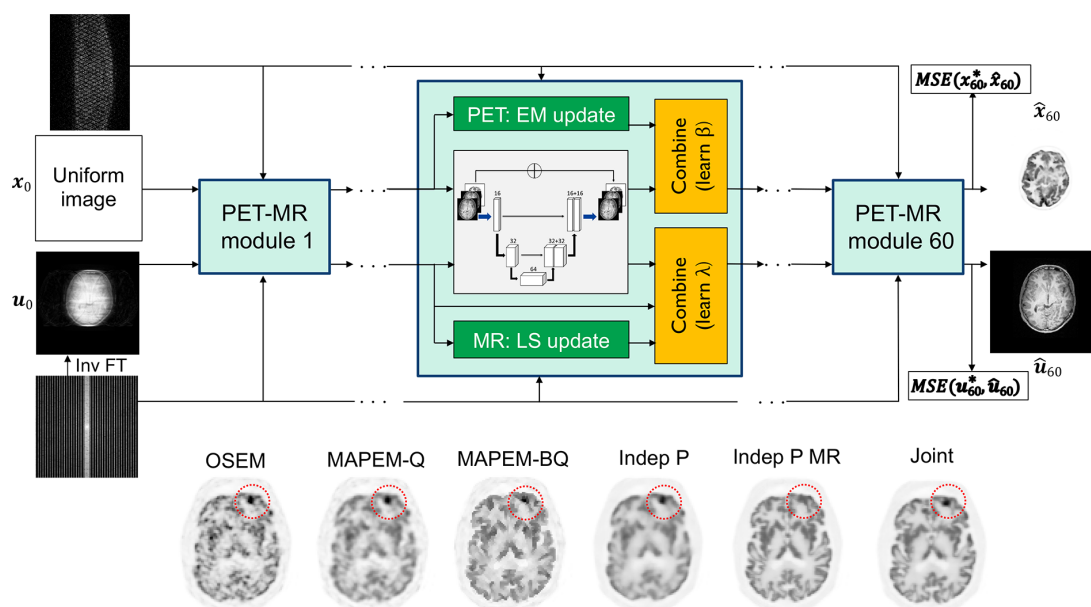
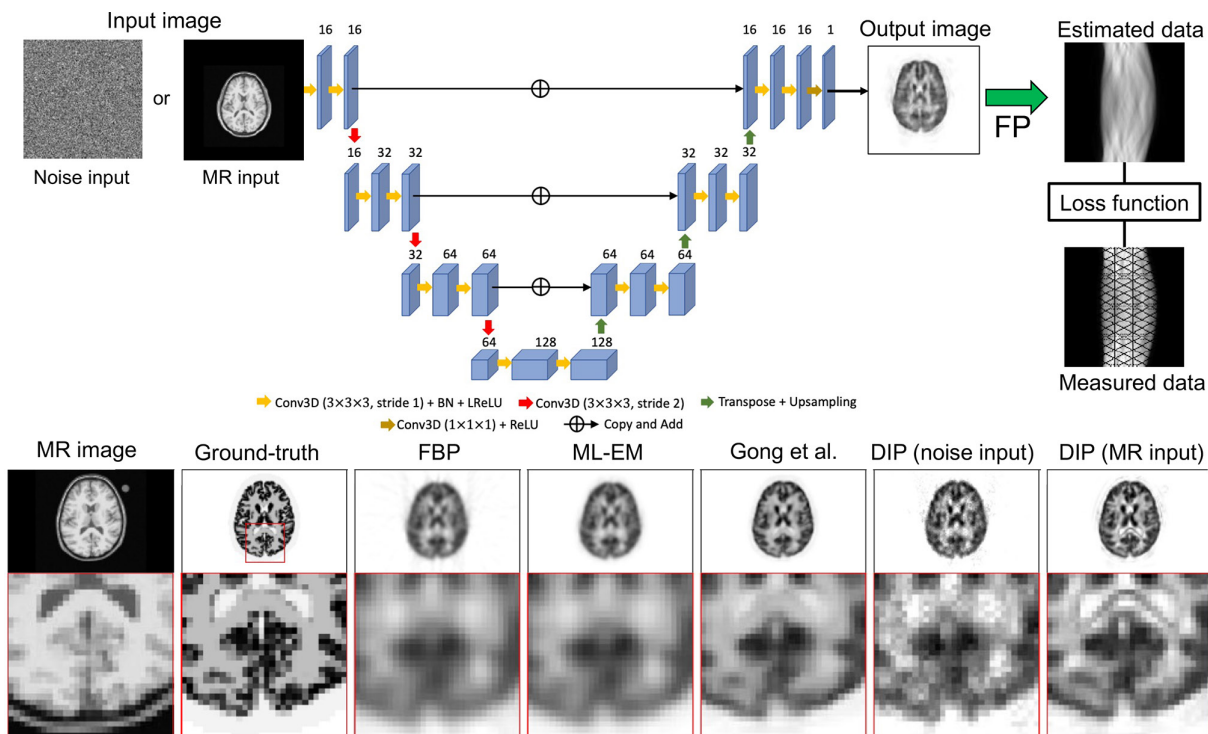


Figure 9. Basic use of the DIP for PET reconstruction, from Hashimoto et al.⁶⁸ FBP, filtered backprojection; DIP, deep image prior; FP, forward projection; PET, positron emission tomography; MLEM, maximum likelihood - expectation maximisation.



resolution PET phantoms⁹⁰ will need much further development, alongside pooling of large real datasets.

A further area of development will be in full exploitation of all types of available data, so that not only supervised learning can be carried out, but also self-supervised learning.⁷⁹ This would allow larger datasets to be used in the absence of high-quality references. Instead of providing high-cost example paired inputs and outputs, only example input data are provided, but crucially with instructions on how to create paired targets automatically for the inputs, hence creating data for supervised learning. This

can be for a pretext task, or indeed directly for a useful end-point task of image reconstruction (as previously shown in Figure 11). Self-supervised approaches with pretext tasks have enabled pretraining of large language models, including powerful pretrained transformer-based generative architectures⁹¹ such as GPT-3⁹² and GPT-4,⁹³ which have now become widely known with significant impact.

One other aspect to note is how AI opens new performance possibilities for PET processing and reconstruction in general. To give just one example area, training data can be used to

Figure 10. Results of using the DIP approach with the kernel method for direct Patlak reconstruction (“DIP recon”), from Gong et al.⁷⁵ DIP, deep image prior; EM, expectation maximisation; KEM, kernel method.

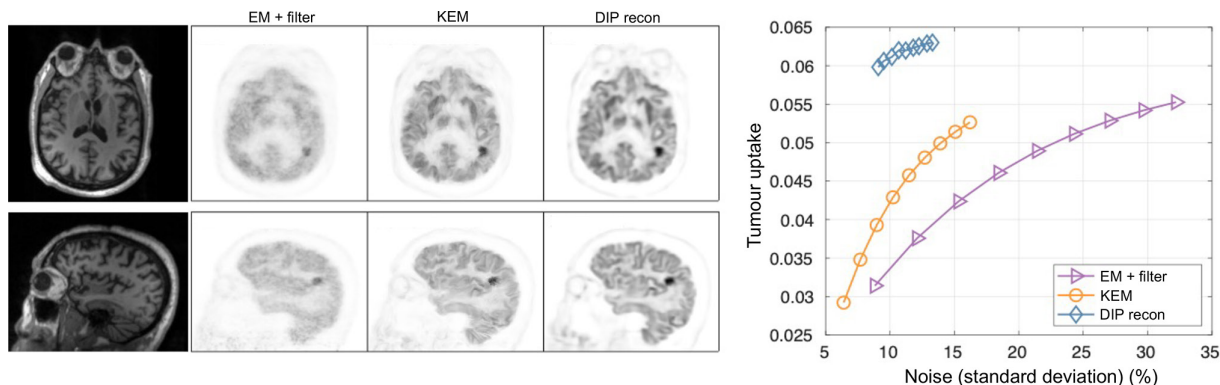


Table 4. DIP and AI-assisted KEM methods for PET image reconstruction.

Name	Architecture [total parameters]	Input	Target	Loss function and optimiser (epochs)
Deep Image Prior Gong et al. 2019 ⁶⁹	U-Net [~2M]	3D anatomical prior (MRI)	An EM iteration in 3D (192 × 192 × 128)	MSE L-BFGS (20 epochs per ADMM iteration, 100 ADMM iterations in total)
Deep Image Prior – with non-negative matrix factorisation Yokota et al. 2019 ⁷⁶	3 U-Nets combined in parallel [~4M]	Noise	Spatial factors—homogeneous tissues with kinetic parameters (128 × 128 × 3)	MSE Adam (20,000)
Deep Image Prior, kernel layer – direct kinetic parametric estimation Gong et al. 2022 ⁷⁵	U-Net [~2M]	T1w 3D MR image	A (dynamic) EM iteration in 3D (128 × 128 × 96 × 7)	MSE L-BFGS (20 epochs per ADMM iteration, 100 ADMM iterations in total)
Neural KEM Li et al. 2022 ⁷⁷	U-Net [~2M]	1-hour composite frames 3D PET images (192 × 192 × 47 × 97)	An EM iteration of coefficient image update	An EM-type surrogate function Adam
Deep Kernel Representation Li et al. 2022 ⁷⁸	U-Net [~2M]	Low-count 3D dynamic PET images	An EM iteration of coefficient images update	MSE [Details not specified]
Deep Image Prior with forward projection Hashimoto et al. 2022 ⁶⁸	U-Net [~2M]	Noise or T1w MR image	2D noisy PET sinogram series (256 × 256 × 64)	L2 loss L-BFGS
Uncertainty-Informed Bayesian Deep Image Prior Sudarshan et al. 2022 ⁷¹	U-Net [details not specified]	3D MR image	An EM iteration in 3D	MSE loss with variance term (BDIP loss) in the first iteration, and BDIP loss + uncertainty-weighted MSE loss in the later iterations Adam
Deep Image Prior with DeepRED Shen et al. 2023 ⁷³	Generator: U-Net [~19M] Denoyer: U-Net [~1M]	Apply transpose of the system matrix to the 2D noisy sinogram (backprojection)	2D whole-body PET image slices at 10M count level (192 × 192)	MSE (sinogram domain) + RED loss (image domain) Adam

2D, two-dimensional; 3D, three-dimensional; AI, artificial intelligence; DIP, deep image prior; BDIP, Bayesian deep image prior; EM, expectation maximisation; KEM, kernel method, PET, positron emission tomography.

Figure 11. Illustration of learning a deep reconstruction operator (a DNN) to operate on sinogram data such that, when backprojected, an image is obtained.⁷⁹ This is an example of self-supervised deep-learned FBP, from Reader,⁷⁹ show just by way of example for the case of a simulated CT slice, but the method is generally applicable. DNN, deep neural network; BP, backprojection; FP, forward projection.

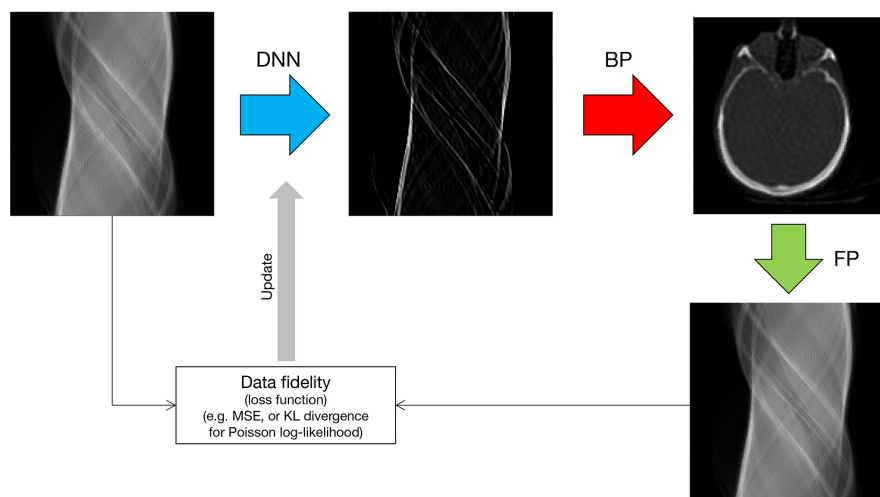
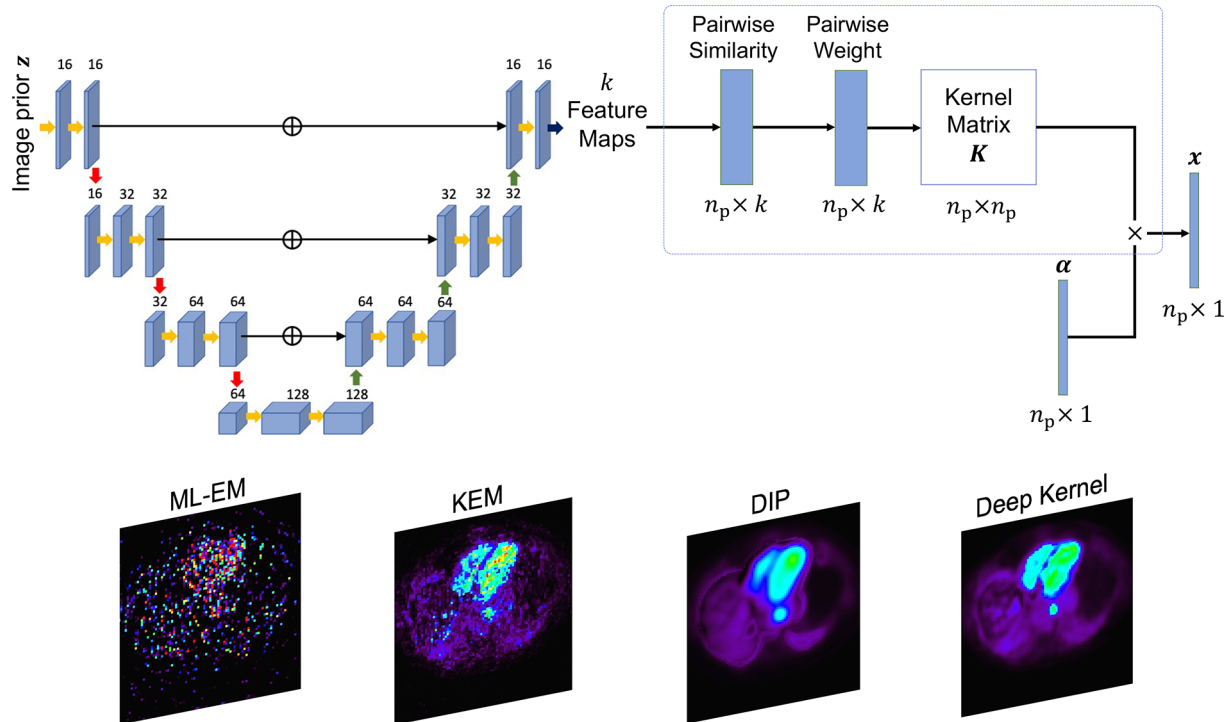


Figure 12. Deep kernel representation by Li and Wang.⁷⁸ A U-Net operates on a prior image (e.g. a 3-frame dynamic image series) to output a set of k different feature maps (a feature vector with k elements for each of the n_p pixels). The similarity of feature vectors between pixels is then used to construct the kernel matrix K , subsequently used in standard KEM. Standard KEM estimates a coefficient vector α , such that the reconstructed image x is found by Kx . The crucial training of the U-Net is based on mapping a 10x count-reduced reconstruction to a full-count reconstruction of the unique data being reconstructed. DIP, deep image prior; KEM, kernel method; MLEM, maximum likelihood–expectation maximisation.



deliver improved performance for separation of multiplexed (e.g. dual tracer^{94,95}) PET image data.^{96,97}

In summary, AI now plays a key role in PET image reconstruction. It has introduced the possibility of making use of high-quality reference data to inform and improve the image reconstruction algorithms that are presently in use. Or, if data external to the patient are unavailable or of concern to use, AI can also provide many methods that only use a patient's own data, potentially alleviating concerns of robustness.

ACKNOWLEDGMENT

This work was supported in part by the Wellcome/EPSRC Centre for Medical Engineering [WT 203148/Z/16/Z], and in part by EPSRC grant number [EP/S032789/1]. For the purposes of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Accepted Author Manuscript version arising, in accordance with King's College London's Rights Retention policy.

REFERENCES

1. Reader AJ, Zaidi H. Advances in PET image reconstruction. *PET Clin* 2007; **2**: 173–90. <https://doi.org/10.1016/j.cpet.2007.08.001>
2. Matej S, Surti S, Jayanthi S, Daube-Witherspoon ME, Lewitt RM, Karp JS. Efficient 3-D TOF PET reconstruction using view-grouped histo-images: DIRECT-direct image reconstruction for TOF. *IEEE Trans Med Imaging* 2009; **28**: 739–51. <https://doi.org/10.1109/TMI.2008.2012034>
3. Barrett HH, White T, Parra LC. List-mode likelihood. *J Opt Soc Am A Opt Image Sci Vis* 1997; **14**: 2914–23. <https://doi.org/10.1364/josaa.14.002914>
4. Reader AJ, Verhaeghe J. 4D image reconstruction for emission tomography. *Phys Med Biol* 2014; **59**: R371–R418. <https://doi.org/10.1088/0031-9155/59/22/R371>
5. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via em algorithm. *J Roy Stat Soc B Met* 1977; **39**: 1–38. <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>
6. Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging* 1982; **1**: 113–22. <https://doi.org/10.1109/TMI.1982.4307558>
7. Lange K, Carson R. EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr* 1984; **8**: 306–16.
8. Qi J, Leahy RM, Hsu C, Farquhar TH, Cherry SR. Fully 3d Bayesian image reconstruction for the ECAT EXACT HR+. *IEEE Trans Nucl Sci* 1998; **45**: 1096–1103. <https://doi.org/10.1109/23.681985>
9. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436–44. <https://doi.org/10.1038/nature14539>
10. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-

- transform manifold learning. *Nature* 2018; **555**: 487–92. <https://doi.org/10.1038/nature25988>
11. Häggström I, Schmidlein 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. <https://doi.org/10.1016/j.media.2019.03.013>
 12. Colsher JG. Fully 3-dimensional positron emission tomography. *Phys Med Biol* 1980; **25**: 103–15. <https://doi.org/10.1088/0031-9155/25/1/010>
 13. Floyd CR. An artificial neural network for SPECT image reconstruction. *IEEE Trans Med Imaging* 1991; **10**: 485–87. <https://doi.org/10.1109/42.97600>
 14. Defrise M, Townsend DW, Clack R. 3-Dimensional image-reconstruction from complete projections. *Phys Med Biol* 1989; **34**: 573–87. <https://doi.org/10.1088/0031-9155/34/5/002>
 15. Belzunce MA, Reader AJ. Time-invariant component-based normalization for a simultaneous PET-MR scanner. *Phys Med Biol* 2016; **61**: 3554–71. <https://doi.org/10.1088/0031-9155/61/9/3554>
 16. Hudson HM, Larkin RS. Accelerated image-reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* 1994; **13**: 601–9. <https://doi.org/10.1109/42.363108>
 17. Reader AJ, Ally S, Bakatselos F, Manavaki R, Walledge RJ, Jeavons AP, et al. One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays. *IEEE Trans Nucl Sci* 2002; **49**: 693–99. <https://doi.org/10.1109/TNS.2002.1039550>
 18. Reader AJ, Julyan PJ, Williams H, Hastings DL, Zweit J. EM algorithm system modeling by image-space techniques for PET reconstruction. *IEEE Trans Nucl Sci* 2003; **50**: 1392–97. <https://doi.org/10.1109/TNS.2003.817327>
 19. Sureau FC, Reader AJ, Comtat C, Leroy C, Ribeiro M-J, Buvat I, et al. Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *J Nucl Med* 2008; **49**: 1000–1008. <https://doi.org/10.2967/jnumed.107.045351>
 20. Vass L, Reader AJ. Synthesized image reconstruction for post-reconstruction resolution recovery. *IEEE Trans Radiat Plasma Med Sci* 2023; **7**: 473–82. <https://doi.org/10.1109/TRPMS.2023.3247489>
 21. Chu G, Tam KC. Three-dimensional imaging in the position camera using fourier techniques. *Phys Med Biol* 1977; **22**: 245–65. <https://doi.org/10.1088/0031-9155/22/2/005>
 22. Erlandsson K, Reader AJ, Flower MA, Ott RJ. A new 3D backprojection and filtering method for pet using all detected events. *IEEE Trans Nucl Sci* 1998; **45**: 1183–88. <https://doi.org/10.1109/23.682000>
 23. Hebert T, Leahy R. A generalized EM algorithm for 3-D bayesian reconstruction from poisson data using Gibbs Priors. *IEEE Trans Med Imaging* 1989; **8**: 194–202. <https://doi.org/10.1109/42.24868>
 24. Levitan E, Herman GT. A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography. *IEEE Trans Med Imaging* 1987; **6**: 185–92. <https://doi.org/10.1109/TMI.1987.4307826>
 25. Wang G, Qi J. Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization. *IEEE Trans Med Imaging* 2012; **31**: 2194–2204. <https://doi.org/10.1109/TMI.2012.2211378>
 26. Ribeiro D, Hallett W, Tavares AAS. Performance evaluation of the Q.Clear reconstruction framework versus conventional reconstruction Algorithms for quantitative brain PET-MR studies. *EJNMMI Phys* 2021; **8**: 41. <https://doi.org/10.1186/s40658-021-00386-3>
 27. Nuyts J, Beque D, Dupont P, Mortelmans L. A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography. *IEEE Trans Nucl Sci* 2002; **49**: 56–60. <https://doi.org/10.1109/TNS.2002.998681>
 28. Bowsher JE, Johnson VE, Turkington TG, Jaszczak RJ, Floyd CR, Coleman RE. Bayesian reconstruction and use of anatomical a priori information for emission tomography. *IEEE Trans Med Imaging* 1996; **15**: 673–86. <https://doi.org/10.1109/42.538945>
 29. Vunckx K, Dupont P, Goffin K, Van Paesschen W, Van Laere K, Nuyts J. Voxel-based comparison of state-of-the-art reconstruction algorithms for 18F-FDG PET brain imaging using simulated and clinical data. *Neuroimage* 2014; **102 Pt 2**: 875–84. <https://doi.org/10.1016/j.neuroimage.2014.06.068>
 30. Mehranian A, Belzunce MA, Niccolini F, Politis M, Prieto C, Turkheimer F, et al. PET image reconstruction using multi-parametric anato-functional priors. *Phys Med Biol* 2017; **62**: 5975–6007. <https://doi.org/10.1088/1361-6560/aa7670>
 31. Belzunce MA, Mehranian A, Reader AJ. Enhancement of partial volume correction in MR-guided PET image reconstruction by using MRI voxel sizes. *IEEE Trans Radiat Plasma Med Sci* 2019; **3**: 315–26. <https://doi.org/10.1109/TRPMS.2018.2881248>
 32. Wang G, Qi J. PET image reconstruction using kernel method. *IEEE Trans Med Imaging* 2015; **34**: 61–71. <https://doi.org/10.1109/TMI.2014.2343916>
 33. Novosad P, Reader AJ. MR-guided dynamic PET reconstruction with the kernel method and spectral temporal basis functions. *Phys Med Biol* 2016; **61**: 4624–44. <https://doi.org/10.1088/0031-9155/61/12/4624>
 34. Bland J, Mehranian A, Belzunce MA, Ellis S, McGinnity CJ, Hammers A, et al. MR-guided kernel EM reconstruction for reduced dose PET imaging. *IEEE Trans Radiat Plasma Med Sci* 2018; **2**: 235–43. <https://doi.org/10.1109/TRPMS.2017.2771490>
 35. De Pierro AR. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Trans Med Imaging* 1995; **14**: 132–37. <https://doi.org/10.1109/42.370409>
 36. Mehranian A, Belzunce MA, McGinnity CJ, Bustin A, Prieto C, Hammers A, et al. Multi-modal synergistic PET and MR reconstruction using mutually weighted quadratic priors. *Magn Reson Med* 2019; **81**: 2120–34. <https://doi.org/10.1002/mrm.27521>
 37. Bland J, Mehranian A, Belzunce MA, Ellis S, da Costa-Luis C, McGinnity CJ, et al. Intercomparison of MR-informed PET image reconstruction methods. *Med Phys* 2019; **46**: 5055–74. <https://doi.org/10.1002/mp.13812>
 38. Reader AJ, Schramm G. Artificial intelligence for PET image reconstruction. *J Nucl Med* 2021; **62**: 1330–33. <https://doi.org/10.2967/jnumed.121.262303>
 39. Hu Z, Xue H, Zhang Q, Gao J, Zhang N, Zou S, et al. DPIR-Net: direct PET image reconstruction based on the Wasserstein generative adversarial network. *IEEE Trans Radiat Plasma Med Sci* 2020; **5**: 35–43. <https://doi.org/10.1109/TRPMS.2020.2995717>
 40. Ma R, Hu J, Sari H, Xue S, Mingels C, Viscione M, et al. An encoder-decoder network for direct image reconstruction on sinograms of a long axial field of view PET. *Eur J Nucl Med Mol Imaging* 2022; **49**: 4464–77. <https://doi.org/10.1007/s00259-022-05861-2>
 41. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017; **39**: 640–51. <https://doi.org/10.1109/TPAMI.2016.2572683>
 42. Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, et al. Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans Med Imaging* 2017; **36**: 2524–35. <https://doi.org/10.1109/TMI.2017.2715284>
 43. Kingma DP, Ba J. Adam: A method for stochastic optimization. 2014.
 44. Jan S, Santin G, Strul D, Staelens S, Assié K, Autret D, et al. GATE: a simulation toolkit

- for PET and SPECT. *Phys Med Biol* 2004; **49**: 4543–61. <https://doi.org/10.1088/0031-9155/49/19/007>
45. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations. 2015.
 46. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets. In: *Advances in Neural Information Processing Systems 27(Nips 2014)*. 2014, pp. 27.
 47. Daube-Witherspoon ME, Muehllehner G. Treatment of axial data in three-dimensional PET. *J Nucl Med* 1987; **28**: 1717–1724.
 48. Whiteley W, Panin V, Zhou C, Cabello J, Bharkhada D, Gregor J. Fastpet: near real-time reconstruction of pet Histo-image data using a neural network. *IEEE Trans Radiat Plasma Med Sci* 2021; **5**: 65–77. <https://doi.org/10.1109/TRPMS.2020.3028364>
 49. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for Biomedical image segmentation. *Med Image Comput Comput Assist Interv, Pt Iii* 2015; **9351**: 234–41. <https://doi.org/10.1007/978-3-319-24574-4>
 50. Feng T, Yao S, Xi C, Zhao Y, Wang R, Wu S, et al. Deep learning-based image reconstruction for TOF PET with DIRECT data partitioning format. *Phys Med Biol* 2021; **66**(16). <https://doi.org/10.1088/1361-6560/ac13fe>
 51. Ote K, Hashimoto F. Deep-learning-based fast TOF-PET image reconstruction using direction information. *Radiol Phys Technol* 2022; **15**: 72–82. <https://doi.org/10.1007/s12194-022-00652-8>
 52. Mehranian A, Reader AJ. Model-based deep learning PET image reconstruction using forward-backward splitting expectation maximisation. *IEEE Trans Radiat Plasma Med Sci* 2020; **5**: 54–64. <https://doi.org/10.1109/TRPMS.2020.3004408>
 53. Reader AJ, Corda G, Mehranian A, Costa-Luis C da, Ellis S, Schnabel JA. Deep learning for PET image reconstruction. *IEEE Trans Radiat Plasma Med Sci* 2021; **5**: 1–25. <https://doi.org/10.1109/TRPMS.2020.3014786>
 54. Gong K, Wu D, Li Q, Kim K, Yang J, El Fakhri G, et al. EMnet: an unrolled deep neural network for PET image reconstruction. *Physics of Medical Imaging*; San Diego, United States. Vol. 10948; 2019. <https://doi.org/10.1117/12.2513096>
 55. 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. <https://doi.org/10.1002/mrm.26977>
 56. Lim H, Chun IY, Dewaraja YK, Fessler JA. Improved low-count quantitative PET reconstruction with an iterative neural network. *IEEE Trans Med Imaging* 2020; **39**: 3512–22. <https://doi.org/10.1109/TMI.2020.2998480>
 57. Li Y, Li J, Liu H, Luo Q, Li X, Gu Y, et al. A Deep ADMM-net for iterative low-count PET reconstruction. *Optics in Health Care and Biomedical Optics X*; Online Only, China. SPIE/COS Photonics Asia: SPIE. Vol. 11553; 2020. <https://doi.org/10.1117/12.2573412>
 58. Adler J, Oktem O. Learned primal-dual reconstruction. *IEEE Trans Med Imaging* 2018; **37**: 1322–32. <https://doi.org/10.1109/TMI.2018.2799231>
 59. Guazzo A, Colarieti-Tosti M. Learned primal dual reconstruction for PET. *J Imaging* 2021; **7**(12): 248. <https://doi.org/10.3390/jimaging7120248>
 60. Ehrhardt MJ, Markiewicz P, Schönlieb C-B. Faster PET reconstruction with non-smooth priors by randomization and preconditioning. *Phys Med Biol* 2019; **64**(22): 225019. <https://doi.org/10.1088/1361-6560/ab3d07>
 61. Gong K, Wu D, Kim K, Yang J, Sun T, El Fakhri G, et al. MAPEM-Net: an unrolled neural network for Fully 3D PET image reconstruction. The Fifteenth International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine; Philadelphia, United States. 2019. <https://doi.org/10.1117/12.2534904>
 62. Hu R, Liu H. TransEM: Residual Swin-Transformer based regularized PET image reconstruction. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022 MICCAI 2022 Lecture Notes in Computer Science 13434*. 2022. <https://doi.org/10.1007/978-3-031-16440-8>
 63. Corda-D'Incan G, Schnabel JA, Reader AJ. Memory-efficient training for fully unrolled deep learned PET image reconstruction with iteration-dependent targets. *IEEE Trans Radiat Plasma Med Sci* 2022; **6**: 552–63. <https://doi.org/10.1109/TRPMS.2021.3101947>
 64. Li Z, Dewaraja YK, Fessler JA. Training end-to-end unrolled iterative neural networks for SPECT image reconstruction. *IEEE Trans Radiat Plasma Med Sci* 2023; **7**: 410–20. <https://doi.org/10.1109/trpms.2023.3240934>
 65. Corda-D'Incan G, Schnabel JA, Hammers A, Reader AJ. Single-modality supervised joint PET-MR image reconstruction. *IEEE Trans Radiat Plasma Med Sci* 2022; **1**. <https://doi.org/10.1109/TRPMS.2023.3283786>
 66. Ulyanov D, Vedaldi A, Lempitsky V. Deep Image Prior. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018. pp. 9446–54.
 67. Singh IRD, Barbano R, Twyman R, et al. Deep Image Prior PET Reconstruction using a SIRF-Based Objective. IEEE Nuclear Science Symposium and Medical Imaging Conference. 2022.
 68. Hashimoto F, Ote K, Onishi Y. PET image reconstruction incorporating deep image prior and a forward projection model. *IEEE Trans Radiat Plasma Med Sci* 2022; **6**: 841–46. <https://doi.org/10.1109/TRPMS.2022.3161569>
 69. Gong K, Catana C, Qi J, Li Q. PET image reconstruction using deep image prior. *IEEE Trans Med Imaging* 2019; **38**: 1655–65. <https://doi.org/10.1109/TMI.2018.2888491>
 70. Cheng Z, Gadelha M, Maji S, Sheldon D. A Bayesian Perspective on the Deep Image Prior. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Long Beach, CA, USA. 2019. pp. 5443–51. <https://doi.org/10.1109/CVPR.2019.00559>
 71. Sudarshan VP, Reddy KPK, Singh M, et al. Uncertainty-Informed Bayesian PET Image Reconstruction Using a Deep Image Prior. ; 2022., pp. 145–55. <https://doi.org/10.1007/978-3-031-17247-2>
 72. Carrillo H, Millardet M, Carlier T, et al. Low-count PET image reconstruction with Bayesian inference over a Deep Prior. SPIE Medical Imaging. Vol. 11596; 2021.
 73. Shen C, Xia W, Ye H, Hou M, Chen H, Liu Y, et al. Unsupervised Bayesian PET reconstruction. *IEEE Trans Radiat Plasma Med Sci* 2023; **7**: 175–90. <https://doi.org/10.1109/TRPMS.2022.3204643>
 74. Metzler CA, Mousavi A, Heckel R et al. (2019) Unsupervised learning with Stein's unbiased risk estimator. BASP Workshops.
 75. Gong K, Catana C, Qi J, Li Q. Direct reconstruction of linear parametric images from dynamic PET using nonlocal deep image prior. *IEEE Trans Med Imaging* 2022; **41**: 680–89. <https://doi.org/10.1109/TMI.2021.3120913>
 76. Yokota T, Kawai K, Sakata M, Kimura Y, Hontani H. Dynamic PET Image Reconstruction Using Nonnegative Matrix Factorization Incorporated With Deep Image Prior. 2019 IEEE/CVF International Conference on Computer Vision (ICCV); Seoul, Korea (South). 2019. <https://doi.org/10.1109/ICCV.2019.00322>
 77. Li S, Gong K, Badawi RD, Kim EJ, Qi J, Wang G. Neural KEM: A kernel method with deep coefficient prior for PET image reconstruction. *IEEE Trans Med Imaging* 2023; **42**: 785–96. <https://doi.org/10.1109/TMI.2022.3217543>

78. Li S, Wang G. Deep kernel representation for image reconstruction in PET. *IEEE Trans Med Imaging* 2022; **41**: 3029–38. <https://doi.org/10.1109/TMI.2022.3176002>
79. Reader AJ. Self-Supervised and Supervised Deep Learning for PET Image Reconstruction. In: *International workshop on machine learning and quantum computing applications in medicine and physics*. 2022.
80. Antun V, Renna F, Poon C, Adcock B, Hansen AC. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc Natl Acad Sci U S A* 2020; **117**: 30088–95. <https://doi.org/10.1073/pnas.1907377117>
81. Luo G, Blumenthal M, Heide M, Uecker M. Bayesian MRI reconstruction with joint uncertainty estimation using diffusion models. *Magn Reson Med* 2023; **90**: 295–311. <https://doi.org/10.1002/mrm.29624>
82. Lakshminarayanan B, Pritzel A, Blundell C. Simple and Scalable predictive uncertainty estimation using deep ensembles. 2016.
83. Gal Y, Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In: Florina BM, Kilian QW, eds. *Proceedings of The 33rd International Conference on Machine Learning*. Proceedings of Machine Learning Research: PMLR. Vol. 48; 2016. pp. 1050–59.
84. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); New Orleans, LA, USA. 2022. <https://doi.org/10.1109/CVPR52688.2022.01042>
85. Hyungjin C, Jeongsol K, Michael Thompson M, et al. Diffusion Posterior Sampling for General Noisy Inverse Problems. The Eleventh International Conference on Learning Representations. 2023.
86. Chung H, Ryu D, McCann MT, et al. Solving 3D Inverse Problems using Pre-trained 2D Diffusion Models. IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.
87. Gong K, Fahrig R, Sabol JM, Yu L. MR guided PET image denoising based on denoising diffusion probabilistic model and data consistency constraint. *Physics of Medical Imaging*; San Diego, United States. SPIE Medical Imaging: SPIE. Vol. 12463; 2023. <https://doi.org/10.1117/12.2653704>
88. Muckley MJ, Riemenschneider B, Radmanesh A, Kim S, Jeong G, Ko J, et al. Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. *IEEE Trans Med Imaging* 2021; **40**: 2306–17. <https://doi.org/10.1109/TMI.2021.3075856>
89. Tolpadi AA, Bharadwaj U, Gao KT, Bhattacharjee R, Gassert FG, Luitjens J, et al. K2S challenge: From undersampled K-space to automatic segmentation. *Bioengineering* 2023; **10**(2): 267. <https://doi.org/10.3390/bioengineering10020267>
90. Belzunce MA, Reader AJ. Technical Note: ultra high-resolution radiotracer-specific digital pet brain phantoms based on the Bigbrain Atlas. *Med Phys* 2020; **47**: 3356–62. <https://doi.org/10.1002/mp.14218>
91. Vaswani A, Shazeer N, Parmar N et al. (2017) Attention is all you need. *Advances in neural information processing systems* 30.
92. Brown T, Mann B, Ryder N, et al. Language Models are Few-Shot Learners. In: Larochelle H, Ranzato M, Hadsell R, et al., eds., Vol. 33. 2020, pp. 1877–1901.
93. Open AI. GPT-4 Technical Report. 2023. <https://doi.org/10.26434/chemrxiv-2023-08774>
94. Kadmas DJ, Rust TC, Hoffman JM. Single-scan dual-tracer FLT+FDG PET tumor characterization. *Phys Med Biol* 2013; **58**: 429–49. <https://doi.org/10.1088/0031-9155/58/3/429>
95. Verhaeghe J, Reader AJ. Simultaneous water activation and glucose metabolic rate imaging with PET. *Phys Med Biol* 2013; **58**: 393–411. <https://doi.org/10.1088/0031-9155/58/3/393>
96. Zeng F, Fang J, Muhashi A, Liu H. Direct reconstruction for simultaneous dual-tracer PET imaging based on multi-task learning. *EJNMMI Res* 2023; **13**(1): 7. <https://doi.org/10.1186/s13550-023-00955-w>
97. Pan B, Marsden PK, Reader AJ. Dual-Tracer PET image separation by deep learning: A simulation study. *Applied Sciences* 2023; **13**: 4089. <https://doi.org/10.3390/app13074089>