

Artificial Intelligence and Deep Learning for Advancing PET Image Reconstruction: State-of-the-Art and Future Directions

Künstliche Intelligenz und Deep Learning für die Weiterentwicklung der PET-Bildrekonstruktion: Stand der Technik und zukünftige Perspektiven



Authors

Dirk Hellwig^{1, 2, 3}, Nils Constantin Hellwig^{1, 3}, Steven Boehner^{1, 2, 3}, Timo Fuchs^{1, 2, 3}, Regina Fischer^{1, 2, 3}, Daniel Schmidt¹

Affiliations

- 1 Department of Nuclear Medicine, University Hospital Regensburg, Regensburg, Germany
- 2 Partner Site Regensburg, Bavarian Center for Cancer Research (BZKF), Regensburg, Germany
- 3 Medical Data Integration Center (MEDIZUKR), University Hospital Regensburg, Regensburg, Germany

Key words

PET image reconstruction, Artificial intelligence, Deep learning, Medical imaging, Quantitative accuracy

received 28.08.2023

accepted after revision 12.10.2023

Bibliography

Nuklearmedizin 2023; 62: 334–342

DOI 10.1055/a-2198-0358

ISSN 0029-5566

© 2023. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence

Prof. Dr. Dirk Hellwig
University Hospital Regensburg, 93042 Regensburg
dirk.hellwig@ukr.de

ABSTRACT

Positron emission tomography (PET) is vital for diagnosing diseases and monitoring treatments. Conventional image reconstruction (IR) techniques like filtered backprojection and iterative algorithms are powerful but face limitations. PET IR can be seen as an image-to-image translation. Artificial intelligence (AI) and deep learning (DL) using multilayer neural networks enable a new approach to this computer vision task. This review aims to provide mutual understanding for nuclear medicine professionals and AI researchers. We outline fundamentals of PET imaging as well as state-of-the-art in AI-based PET IR with its typical algorithms and DL architectures. Advances improve resolution and contrast recovery, reduce noise, and remove artifacts via inferred attenuation and scatter correction, sinogram inpainting, denoising, and super-resolution refinement. Kernel-priors support list-mode reconstruction, motion correction, and parametric imaging. Hybrid approaches combine AI with conventional IR. Challenges of AI-assisted PET IR include availability of training data, cross-scanner compatibility, and the risk of hallucinated lesions. The need for rigorous evaluations, including quantitative phantom validation and visual comparison of diagnostic accuracy against conventional IR, is highlighted along with regulatory issues. First approved AI-based applications are clinically available, and its impact is foreseeable. Emerging trends, such as the integration of multimodal imaging and the use of data from previous imaging visits, highlight future potentials. Continued collaborative research promises significant improvements in image quality, quantitative accuracy, and diagnostic performance, ultimately leading to the integration of AI-based IR into routine PET imaging protocols.

Objective and Scope of the Review

The objective of this comprehensive review is to provide a broad and understandable overview of the current advancements, chal-

lenges, and promising avenues in integrating artificial intelligence (AI) techniques into the field of positron emission tomography (PET) image reconstruction (IR). This review is intended for nuclear medicine physicians, medical physicists, nuclear medicine tech-

nicians and researchers, as well as AI researchers interested in such applications, and bridges the gap between these fields. By exploring the intersection of PET IR and AI, this review aims to provide readers with a clear understanding of how AI-driven approaches are reshaping the art of PET imaging, enhancing image quality and diagnostic performance.

The review deals with the IR process, but not with AI-driven analysis of previously reconstructed PET images, which is covered in the companion articles in this journal issue. It highlights the potential clinical impact of AI techniques and the ongoing collaborative efforts needed to successfully integrate these innovative methods into routine PET imaging protocols. Through this broad scope, the review aims to facilitate meaningful dialogue and inspire further collaboration between the nuclear medicine and AI communities to foster advances that hold promise for the future of PET imaging. The search strategy for relevant publications included an exploration of literature sources spanning from historical developments and foundational mathematical principles to contemporary advancements in AI-assisted PET IR to ensure a comprehensive coverage of the topic.

Fundamentals of PET Imaging

PET imaging has revolutionized medical imaging by providing valuable insights into the metabolic processes and targeting diseases occurring within the human body. It provides less anatomical information than computed tomography (CT) or magnetic resonance imaging (MRI), so it is usually combined as PET/CT or PET/MR with image fusion of images from both modalities from a single hybrid scanner.

PET imaging involves the use of radiopharmaceuticals that emit positrons, which subsequently annihilate with electrons as their anti-particles. This results in the simultaneous nearly diametral emission of two photons which are detected by a ring of detectors surrounding the patient. The radiation transport is affected by absorption and scattering by the surrounding material causing loss of energy and deviated pathway of photons. The annihilation photons are detected in coincidence on lines-of-response between detector pairs and counted by the PET scanner. Small time differences between the detection of the photons are used for the time-of-flight (TOF) technique to limit the site of origin of the photons and thus improve the signal-to-noise ratio. Measurement effects such as dead time and random coincidences are corrected. Since radioactive decay follows Poisson statistics, with lower activity the random noise increases and degrades image quality. A broad spectrum of approved PET tracers is clinically available to diagnose and monitor many oncologic and non-oncologic diseases. Mostly fluorine-18 or gallium-68 serve as the radioactive label for the metabolic substrates or for the ligands to target structures. The radioactivity of the PET tracers causes radiation exposure to the patient and personnel, requiring the dose to be minimized.

The emission data acquired during a PET acquisition is usually sorted into sinograms (a sorted collection of projection data acquired at various angles) of cumulated events or in list-mode as time-stamped events which need resorting prior to further use.

Reconstructed images finally depict the distribution of radiotracer concentrations in the body which varies over time as determined by biochemical and physiological conditions. IR is a critical step in PET imaging, aiming to transform the acquired raw data into accurate and high-quality images with adequate spatial and temporal resolution for diagnostic interpretation [1].

Conventional PET IR: Filtered Backprojection and Iterative Algorithms

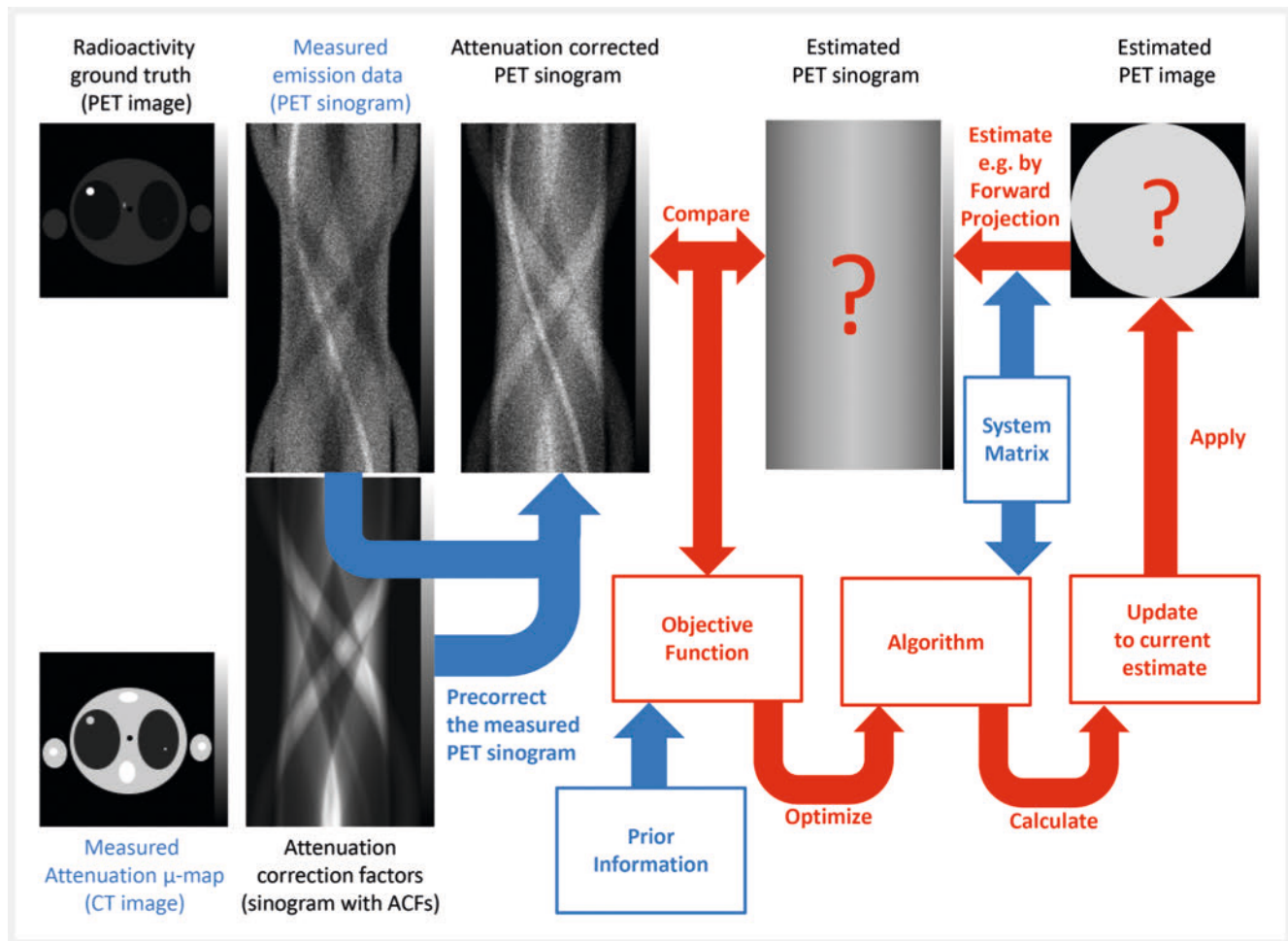
Conventional PET IR methods, such as filtered backprojection (FBP) and iterative reconstruction algorithms, have been widely used for several decades. For FBP the IR problem is considered as a Radon transformation of the image plane to a sinogram [2] which is solved by use of the central slice theorem using Fourier transformations [3].

FBP reconstructs images by backprojecting the detected radiation events from sinograms onto the grid of the image space using filters with precalculated weights analytically derived from ideal noise-free projection conditions. Empirically determined filters cut off higher spatial frequencies to limit noise. Despite its fast calculation, this approach suffers severely from the amplification of low-count related noise and from artifacts in high-contrast imaging situations, resulting in limited spatial resolution and reduced lesion detection [1].

Iterative image reconstruction (IIR) methods interpret the reconstruction problem as a system of algebraic equations in which the vector of a voxelized radioactivity distribution is multiplied by a (precalculated) transition matrix based on the imaging physics, the so-called system matrix. It models physical effects such as the properties of the decay, radiation transport and detection processes including interaction of radiation within the patient, detector geometry, detection efficiency, and inherent resolution of the PET scanner. The matrix product corresponds to the expected projection data. ► **Fig. 1** comprehends the main steps of an IIR algorithm to reconstruct PET images from a PET/CT acquisition.

In common, the reconstruction task to invert the linear mapping from the given radioactivity distribution (as ground truth) to the measured detector events (as observation) is an ill-posed problem. The inversion of the system matrix is practically impossible, not only due to photon count limitations caused by the natural Poisson statistics of radioactive decays, but also from the high number of matrix elements with technically unavoidable uncertainties, e. g. of detector efficiencies which can only be measured with a statistical error. The inversion of such linear mappings can be achieved by FBP. For the sake of completeness, the Moore-Penrose pseudo-inverse should be mentioned here as a mathematical tool to find approximate solutions to ill-posed linear problems, namely for non-square and singular matrices, that don't have a true inverse. It is also noise affected so that its calculation by means of singular value decomposition (SVD) needs an empirically determined limitation of the singular value spectrum for noise suppression [4], similar to the filter parameters in FBP reconstructions.

IIR algorithms aim to estimate the radioactivity distribution that best fits the measured emission data with respect to the sta-



► **Fig. 1** Schematic data flow and workflow of conventional iterative image reconstruction, illustrated with a simulated radioactivity distribution and attenuation maps of an exemplary axial slice of PET/CT in lung cancer. Measured PET sinograms are noisy due to the Poisson statistics of radioactive decay. The CT image corresponds to the attenuation map needed to pre-correct the measured data in the PET sinogram. The System Matrix describes the physical conditions of the PET scanner like geometry and detector properties. Prior knowledge and measured data (colored blue) are used for data preparation and the iterative reconstruction process. The current estimate of the PET image is converted into an estimate of the corresponding PET sinogram using the system matrix. To compare this estimate with the measured data, an objective function is used, which may be motivated by prior knowledge such as constraints on the expected image resolution. To optimize this objective function, specific algorithms compute updates to obtain a better estimate of the PET image. The iterative process (colored red) can be terminated after a predefined number of cycles or when the objective function meets the quality criteria.

tistical properties of the Poisson-distributed raw data from radioactive decays. They often require many iterations leading to high computational effort. IIR has been used for oncology PET imaging since the early 1990s [5]. To reduce the computer workload, methods for accelerating convergence were introduced, e. g. Ordered Subset Expectation Maximization (OSEM) [6]. As a multi-million parameter estimation, IIR tends to overfit with high amplitude patterns ("checkerboard effect" or "night sky artefact") as the number of iterations increases [7]. As an effect of modelling the Point Spread Function (PSF) of the PET system, Gibbs artifacts may occur as overestimation of radioactivity levels in small structures, indicating the need for careful consideration when employing IIR algorithms for quantitative PET data analysis [8].

Regularized IR methods, like Maximum a Posteriori (MAP) algorithms, combine image likelihood and prior probabilities [9]. The prior functions as a penalty term to suppress artifacts and im-

prove IIR accuracy by enforcing image smoothness. The feasibility of anatomical priors was proven in earlier years [10] and research continues with MR-guided kernels for IIR of reduced dose PET imaging [11].

Despite all achievements, conventional IIR techniques struggle with noise amplification, limited spatial resolution, and substantial computational demands.

Fundamentals of AI, Machine Learning and Deep Learning

The emergence of AI has inaugurated a novel era in medical imaging, where data-driven techniques are being employed to enhance diagnostics and patient care. Machine Learning (ML) is a specialized field within AI, focusing on the development of algo-

gorithms and models that enable systems to learn patterns and make predictions or decisions from data without being explicitly programmed [12]. Deep Learning (DL), a sub-field of ML, is characterized by its ability to automatically learn hierarchical features and complex patterns from large amounts of data by means of neural networks (NNs).

NNs are computational models inspired by the human brain's neural connections. They consist of interconnected nodes, or "neurons" to simulate nervous activity [13], organized in layers. Neurons have one or more inputs. Their weighted sum serves as the input for mostly non-linear activation functions, which then generate the neurons' output. NNs may consist of multiple layers as a cascade of mappings, with hidden layers between input and output layers, building deep NNs. According to the universal approximation theorem (UAT), multilayer feedforward networks are universal approximators [14]. By using the non-linearity between layers, many practical and useful mappings can be well approximated if enough layers are used. Introduced to AI to simulate human-like learning, NNs excel in capturing complex patterns from data through a process called training.

During the training phase, NNs learn by adjusting their internal parameters using a designated training dataset. The training is an iterative process that involves updating the network's parameters to minimize the discrepancy between predicted and actual outputs (the ground truth), quantified by a loss function. This is typically achieved by using optimization algorithms like Adam [15], which adjusts the weights and biases of the network's layers based on the gradients of the loss function. Hyperparameters, such as learning rate and batch size, significantly influence the rate and stability of convergence in this process. Regularization techniques, like dropout or L2 regularization, are commonly applied to prevent overfitting, an undesirable ML behavior that occurs when the ML model provides accurate predictions for training data but not for new data [14]. Additionally, data augmentation is used to increase the effective size of the training dataset. Proper validation techniques and monitoring of training curves are essential to avoid overfitting and ensure the generalizability of the trained model [16].

To prevent overfitting and fine-tune the model's hyperparameters, a separate validation dataset is utilized. This validation dataset helps in assessing the model's performance on new, unseen data and aids in making decisions about architecture and hyperparameters. Once the training is complete, the model's final evaluation is conducted using an entirely independent test dataset, which ensures an objective assessment of the NN's generalization capability. For first own impressions, the interested reader may have a look at <https://playground.tensorflow.org>.

Learning takes time and resources. The concept of transfer learning has facilitated the transfer of knowledge gained from one task to another. In NNs, transfer learning is accomplished by using a model pretrained on a data-intensive task and then fine-tuning it on the target task with a smaller dataset to adapt the learned features. Thereby, training is accelerated and the generalization capabilities of AI models are enhanced [17]. Once trained, NN-based models can provide fast inference, making them suitable for real-time applications.

Architectures of Deep Neural Networks

Some specific DL architectures have emerged as powerful tools, revolutionizing the field's capabilities and potential. Images are too big for normal NNs as the full connection of neurons sum up to tens and hundreds of million weights. Here, we delve into three prominent architectures: Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), and U-Net architectures, exploring their unique characteristics. These architectures, coupled with breakthroughs in optimization techniques, like Adam [15] and RMSprop [18], have revolutionized the speed and precision with of AI-driven image analysis.

Convolutional Neural Networks

CNNs are particularly adept at image analysis due to their inherent ability to automatically learn spatial, hierarchical (and translation-invariant) features from data [19]. Instead of general matrix multiplications of fully connected NNs, they employ convolutional layers elementwise multiplication of kernel weights and corresponding values in the image to extract spatial patterns and hierarchical representations. CNNs are highly effective across diverse medical imaging tasks, from image transformation and segmentation to disease classification [16], which are also essential for accurate IR.

The use of convolution kernels in CNNs is the key to translation-invariant feature extraction, resulting in a reduction of overall network parameters and thereby also reducing computational effort and overfitting. The concept of a kernel has a significant connection to ML, particularly in the context of support vector machines (SVMs), which are used in classification and regression tasks in high-dimensional spaces [20]. In ML, a kernel is a function that computes the similarity or inner product between two data points in a higher-dimensional feature space, without explicitly transforming the data into that space. The connection between the usage of filter kernels in CNNs and the kernel trick of SVMs is that with both techniques complex patterns are learned and captured more efficiently by transforming the data into higher dimensional spaces.

Kernel weights dictate the features extracted in each channel. Although the specifics of learned kernels are concealed within a black box, they correlate with the primate visual cortex's neural wiring and vision physiology [21]. These kernels may resemble Gabor functions, well-suited for detecting image textures [22]. CNNs frequently employ pooling steps to contract adjacent pixels, achieving dimensional reduction. For a firsthand experience of CNN performance in classification tasks, interested readers can explore: <https://poloclub.github.io/cnn-explainer/>.

U-Nets

U-Net architectures are a class of CNNs specifically designed for semantic segmentation tasks. They derive their name from their characteristic U-shaped structure, consisting of an encoder path for feature extraction and a decoder path for feature localization. U-Nets enable pixelwise classification by preserving spatial information through skip connections that connect the encoder and decoder layers. This makes them particularly effective for tasks like tumor

segmentation, where precise delineation of structures is crucial. U-Nets have demonstrated impressive performance [23].

Generative Adversarial Networks

GANs offer a unique approach to data generation. Consisting of two networks, a generator and a discriminator, GANs work in tandem to create highly realistic synthetic data. The generator learns to produce images that resemble real data by deconvoluting a noise pattern, while the discriminator assesses the authenticity of these generated samples. Through adversarial training, GANs can produce high-fidelity images that are virtually indistinguishable from real ones [24]. By learning the distribution of given images, such a model enables the synthesis of new images from this distribution. Conditional GANs (cGANs) allow the generation of images using preset conditions, e. g. with desired disease foci or contours of structures to be rendered. In medical imaging, GANs are used for tasks like data augmentation to enhance the quality and diversity of medical image datasets, super-resolution enhancement or denoising, and image-to-image translation, even in cross-modality image synthesis. However, with synthetic images of GANs, it is important to remember that a false disease state may be represented.

Concepts to Apply DL to PET IR

DL has emerged as promising approach for enhancing PET IR. The IR process can be interpreted as an image-to-image transformation, since a set of measured sinograms as input is used to infer the underlying volumetric radioactivity distribution as target. The learning process can be interpreted as finding a mapping between the measured sinogram and the underlying ground truth of given radioactivity distribution.

Direct DL Methods

Direct DL methods for PET IR bypass the need for models of the imaging system and statistical noise, instead focusing on large-scale training data to establish a full end-to-end mapping from raw sinogram data directly to reconstructed images. While this approach may avoid potential modeling errors, it neglects the expertise and progress gained from years of model-based reconstruction development. Additionally, the exclusion of these models might lead to inexplicable mappings, potentially affecting confidence, particularly for unforeseen inputs. However, notable examples of direct DL methods include AUTOMAP [25], DeepPET [26], Liu et al. (using cGANs) [27], and DirectPET [28], although most applications have been limited to small 2-D slices.

Direct: Fully Connected Layers with CNNs

AUTOMAP utilizes fully connected layers to learn a mapping akin to the inverse Radon transform, followed by a CNN for denoising. Although primarily designed for MR image reconstruction, AUTOMAP was also adapted for PET reconstruction [25]. However, PET reconstruction results using single slice rebinned input sinograms were less compelling, exhibiting lower visual quality than conventional Poisson OSEM [29]. The approach in DirectPET [28] to use

volumes with 16 slices outperformed the vendor's standard IRR with respect to signal-to-noise ratio.

Direct: Convolutional Encoder–Decoder

DeepPET [26] applies the encoder-latent space-decoder architecture for direct PET IR. The convolutional encoder-decoder (CED) method employs convolutional downsampling to progress from sinograms to a feature-rich latent space representation, which is then upsampled in the decoder to yield the PET image. This eliminates the need for modeling assumptions by learning the imaging physics and noise distribution, resulting in significantly accelerated image reconstruction and increased adaptability to real data. However, while simulations show promising results, real data, especially brain data, still requires further refinement due to the absence of high-quality reference data [14]. An adversarial variant of this concept, proposed by Liu et al. [27], replaces CED with a U-Net conditional generator and adds a discriminator network, and an extended version with a discriminator was later suggested by Hu et al. [30].

Ma et al. report the prototypic implementation of an encoder-decoder network, based on the VGG19 network pre-trained on the ImageNet database, for direct IR on sinograms of a long axial field of view PET [31]. They demonstrated the potential of DL to learn complex IR principles such as projection, normalization, attenuation correction, and scattering correction by training with real clinical data. However, the prototype failed to accurately reconstruct PET scans of a physical phantom or in cases with extreme anatomy [31].

DL Regularization Methods

Regularization is used to reduce or minimize noise in the context of image processing or data analysis, often by adding a penalty term to the model's objective function, to find a simpler and smoother solution that generalizes better and more robust to new, unseen data. DL regularization methods mainly work in the image domain. They use pre-existing models and potential functions for regularization penalties. Unlike conventional model-based techniques that used handcrafted priors (typically Gaussian shaped priors motivated by the spatial resolution of the PET scanner), these approaches integrate data-driven priors into image reconstruction, enhancing the regularization component of the process while retaining standard imaging physics and statistics. For practical implementation, so called 'generators' can be employed for some crucial steps within a conventional framework of iterative reconstruction. To optimize the reconstruction objective function, these generators may serve 1. for synthesis purposes, e. g. as denoisers or conditional generators, or 2. for analysis purposes, e. g. by creating a sparse coded description of image features which corresponds mathematically to a dimensional domain-transformation with evaluation in so-called latent spaces.

Deep Learning for Image Generation: Synthesis Regularization

In synthesis-based regularization, deep learning functions as an image constraint, utilizing sophisticated deep mapping generators to create image estimates. This concept demands that image

estimates align with the output of a deep network operating on input code vectors. The elements of such an input code vector may be the intensities within a predefined partitioning of the image domain, e. g. within predefined regions-of-interest covering the whole scan area. This approach offers flexibility in integrating advanced denoising techniques into PET image reconstruction, even within fully 3D contexts. Within this framework, three primary approaches have emerged:

- Estimating an input code vector for a fixed deep network to optimize the reconstruction objective function
- Using a variable input code while estimating network parameters for optimization
- Combined estimation of both network parameters and input code vectors, either simultaneously or alternately.

To exemplify, Gong et al. [32] introduced a method that seamlessly integrates conventional IR techniques into the broader algorithmic framework of the alternating direction method of multipliers (ADMM). Their work initially addresses a conventional MAP-EM problem for image update, employing a quadratic penalty with a prior image generated from a convolutional neural network (CNN) operating on the current code vector estimate. An extended approach employs a fixed input vector, inspired by the "deep image prior" (DIP) [33], where a deep CNN is trained to map the fixed prior to match the current MAP-EM update, with a particular focus on regions where CNN-based reconstruction had previously struggled. Notably, this method is unsupervised, requiring no training data, and has been compared to CNN penalty methods. Moreover, the methodology has been extended to 4D IR and dynamic PET, further demonstrating its versatility and potential for enhancing parametric IR [34, 35] as well as for motion correction [36].

Deep Learning for Analysis Regularization

In the context of analysis-based regularization, deep networks are seamlessly integrated into conventional prior or penalty functions, constituting a fundamental element of IR through an analysis regularization strategy [37]. Rather than imposing stringent constraints as in synthesis strategies, such as image generation or denoising, analysis priors aim for reconstructed images to align with measured data (e. g., Poisson log likelihood) while adhering to the proximity of a deep denoised image version. This nuanced approach resembles the contrast between MAP-EM and KEM methods [38], with analysis-based techniques offering a less restrictive framework.

Deep Learning for the Entire Prior: Unfolded Methods

The concept of physics-informed DL involves merging the strength of AI with our existing comprehension of imaging physics and statistical models, deploying AI specifically for the aspects of reconstruction where confidence is lacking, such as the precise regularization method and its strength. This strategy offers interpretability, a critical aspect in clinical imaging, while replacing the potential function with deep-learned mappings through the unfolding of conventional IR steps [39, 40]. In this approach, DL covers the complete penalty or prior, eliminating the requirement for explicit analytic, intuitive, or handcrafted components. The IR al-

gorithms are unrolled, transforming them into a cascade of blocks, where each block represents an iterative update and can be explicitly defined as a processing operator [40, 41]. These blocks, combining trainable gradient-based components for the penalty and fixed operator components for data consistency, establish a deep network that integrates partial or complete reconstruction operators with deep denoising operators. A key feature is the use of a deep-learned denoised prior image from the previous reconstruction estimate, creating a recurring loop [40, 41].

Three major unrolled methods emerged for PET IR with integrated DL for regularization: BCD-Net [39], MAPEM-Net [40], and FBSEM-Net [41]. BCD-Net's training focuses on the block level, denoising updates to align with a high-quality reference [39]. In contrast, both MAPEM-Net and FBSEM-Net perform training based on the final iteration, necessitating backpropagation through all blocks during training for parameter updates [40, 41]. MAPEM-Net conducts two MAP-EM updates per block, aligning the final iteration with a high-quality reference [40]. For BCD-Net, training is executed for individual block-dependent denoisers, matching the iteration's outcome with the high-quality reference, thus avoiding backpropagation across all blocks [39]. In FBSEM-Net, training ensures that the last iteration matches a high-quality reference, leveraging a fixed image like MRI in an L2 norm penalty for the MAP-EM update [41]. Notably, FBSEM demonstrates an ability to mitigate PSF Gibbs artifacts.

Cordea-D'Incan et al. present an innovative approach for joint PET-MR IR [42], unrolling the MAP-EM algorithm for PET and the Landweber algorithm for MR through a DL joint regularization step. Their investigation of loss function selection demonstrates that a network trained with a single-modality loss achieves superior global reconstruction accuracy for PET and improved PET-specific feature reconstruction, while joint reconstruction gains for MR are primarily observed with highly undersampled data, showcasing potential benefits of the proposed framework for multimodal IR.

Deep Learning for Preprocessing and Post-Processing

Preprocessing of raw sinograms is of interest for the correction of detector failures. Whiteley et al. used an inpainting technique to successfully correct for count loss due to the failure of detector blocks [43].

Both low-dose PET data and images can be upgraded post-reconstruction to their high-dose counterparts via DL denoising approaches to enhance spatial resolution and mitigate noise [44, 45, 46, 47]. An alternative approach involves backprojected images, where raw PET data (sinograms or list-mode data) are initially backprojected into a 3D image array before undergoing reconstruction to restore the quantitative radiotracer distribution [48]. Exploiting backprojected images, including time-of-flight information for the so called histo-images, has shown promise for deep-learned mappings [49].

Another important step of post-reconstruction processing is attenuation and scatter correction. Shiri et al. demonstrated the feasibility of direct attenuation and scatter correction of whole-

body [^{18}F]-FDG PET images using emission-only data via a deep residual network. The proposed approach achieved accurate attenuation and scatter correction without the need for anatomical images, such as CT and MRI [50].

Li et al. [51] addressed motion artifacts in PET imaging caused by respiratory movements by an unsupervised non-rigid image registration framework based on DL. This hybrid methodology combines deep NNs for image warping and deformation field refinement with iterative IR for motion-compensated PET image generation. The study demonstrates the potential of this hybrid approach to enhance image quality, lesion contrast, and boundary sharpness, while offering improved performance compared to traditional iterative registration methods in the context of respiratory-gated PET imaging.

AI-based super-resolution techniques [52] offer a powerful approach to enhance the spatial resolution of PET images, improving image quality and enabling more accurate and detailed analysis. By leveraging DL models trained on large datasets of images or image patches [53], these techniques can recover fine details, preserve quantitative accuracy, address noise-related challenges, and impact various clinical applications, including small lesion detection, anatomical localization, and quantitative analysis. The integration of AI-based superresolution in PET imaging has the potential to advance diagnostic capabilities and improve patient care.

Challenges and Future Directions

The application of AI to PET IR holds immense promise, yet several challenges must be navigated to ensure successful integration into clinical practice.

Computational Efficiency and Real-Time Reconstruction

One of the foremost challenges lies in the computational demands of DL algorithms utilized in PET IR. The efficiency of these algorithms, both during training and inference, is crucial for clinical feasibility. Real-time or near-real-time reconstruction is desirable for prompt clinical decision-making. This necessitates the development of efficient algorithms and optimized implementation strategies to accommodate clinical timelines [54].

Availability and Representation of Training Data

AI models heavily rely on large, diverse, and representative datasets for effective learning and generalization. Acquiring such datasets is resource-intensive and requires meticulous collation of PET data from various sources, considering different scanners, acquisition protocols, anatomical regions, and potential sources of variability [16, 54].

Interpretability and Trust

The inherent complexity of DL models often leads to them being treated as black boxes, making it challenging to decipher the reasoning behind their decisions. This lack of interpretability can hinder the

trust and acceptance of AI-driven reconstruction methods in clinical settings, where explainability and transparency are vital [55].

Clinical Validation and Translation

A critical direction for the future is the clinical validation and translation of AI-driven PET image reconstruction techniques. Rigorous studies and clinical trials across multiple centers are needed to establish the clinical utility, generalizability, and impact of these methods on diverse patient populations, imaging protocols, and disease conditions [54].

Regulatory and Ethical Considerations

The integration of AI-based techniques into clinical practice brings forth regulatory and ethical challenges. Regulatory frameworks such as the Medical Device Regulation (MDR) in the European Union demand stringent validation, safety, and efficacy standards for AI-based medical devices. Compliance with these regulations, alongside ethical considerations including data privacy and patient consent, is imperative when working with large datasets [56].

Pathways for Future Development

To navigate these challenges, a collaborative approach among researchers, clinicians, and regulatory bodies is indispensable. Initiatives like data sharing, multi-center collaborations, transfer learning, standardized algorithms, and dedicated regulatory frameworks can collectively address these obstacles and pave the way for the seamless integration of AI-assisted PET image reconstruction into routine imaging protocols [57].

In summary, while the journey toward integrating AI into PET imaging workflows comes with significant challenges, the potential for enhancing image quality, diagnostic accuracy, and patient care is undeniable. By addressing these challenges collectively and advancing research, the future promises the harmonious merger of AI-driven PET image reconstruction with routine clinical practice.

Funding

Bavarian Center for Cancer Research (BZKF) (ZB-001) | Bundesministerium für Bildung und Forschung (ABIDE_MI: 01ZZ2061 T, DIFUTURE: 01ZZ1804H, NUM: 01KX2121) | <http://dx.doi.org/10.13039/501100002347>

Conflict of Interest

The authors declare that they have no conflict of interest.

Literature

- [1] Hoeschen C, Rafecas M, Aspelmeier T. Algorithms for Image Reconstruction. In: Cantone MC, Hoeschen C, (eds.) *Radiation Physics for Nuclear Medicine* Berlin, Heidelberg: Springer Berlin Heidelberg; 2011: 211–232
- [2] Radon J. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte Über Verhandlungen K-Sächs Ges Wiss Zu Leipz Math-Phys Kl* 1917; 69: 262–277
- [3] Bracewell R. Strip Integration in Radio Astronomy. *Aust J Phys* 1956; 9: 198. doi:10.1071/PH560198

- [4] López-Montes A, Galve P, Udías JM et al. Real-Time 3 D PET Image with Pseudoinverse Reconstruction. *Appl Sci* 2020; 10: 2829. doi:10.3390/app10082829
- [5] Bares R, Klever P, Hellwig D et al. Pancreatic cancer detected by positron emission tomography with 18F-labelled deoxyglucose: method and first results. *Nucl Med Commun* 1993; 14: 596–601. doi:10.1097/00006231-199307000-00013
- [6] Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* 1994; 13: 601–609. doi:10.1109/42.363108
- [7] Levitan E, Herman GT. A Maximum a Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography. *IEEE Trans Med Imaging* 1987; 6: 185–192. doi:10.1109/TMI.1987.4307826
- [8] Rogasch JM, Hofheinz F, Lougovski A et al. The influence of different signal-to-background ratios on spatial resolution and F18-FDG-PET quantification using point spread function and time-of-flight reconstruction. *EJNMMI Phys* 2014; 1: 12. doi:10.1186/2197-7364-1-12
- [9] Lange K, Bahn M, Little R. A Theoretical Study of Some Maximum Likelihood Algorithms for Emission and Transmission Tomography. *IEEE Trans Med Imaging* 1987; 6: 106–114. doi:10.1109/TMI.1987.4307810
- [10] Lipinski B, Herzog H, Rota Kops E et al. Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information. *IEEE Trans Med Imaging* 1997; 16: 129–136. doi:10.1109/42.563658
- [11] Bland J, Mehranian A, Belzunce MA et al. MR-Guided Kernel EM Reconstruction for Reduced Dose PET Imaging. *IEEE Trans Radiat Plasma Med Sci* 2018; 2: 235–243. doi:10.1109/TRPMS.2017.2771490
- [12] Bishop CM. *Pattern Recognition and Machine Learning*. Softcover reprint of the original 1st edition 2006 (corrected at 8th printing 2009). New York, NY: Springer New York; 2016
- [13] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 1943; 5: 115–133. doi:10.1007/BF02478259
- [14] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Netw* 1989; 2: 359–366. doi:10.1016/0893-6080(89)90020-8
- [15] Kingma DP, Ba J. Adam: A Method for Stochastic Optimization; 2017. doi:10.48550/arXiv.1412.6980
- [16] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; 521: 436–444. doi:10.1038/nature14539
- [17] He K, Zhang X, Ren S et al. Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE; 2016: 770–778. doi:10.1109/CVPR.2016.90
- [18] Tieleman T, Hinton G. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA Neural Netw Mach Learn* 2012; 4: 26–31
- [19] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2017; 60: 84–90. doi:10.1145/3065386
- [20] Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: *Proceedings of the fifth annual workshop on Computational learning theory*. Pittsburgh Pennsylvania USA: ACM; 1992: 144–152. doi:10.1145/130385.130401
- [21] De A, Horwitz GD. Spatial receptive field structure of double-opponent cells in macaque V1. *J Neurophysiol* 2021; 125: 843–857. doi:10.1152/jn.00547.2020
- [22] Turner MR. Texture discrimination by Gabor functions. *Biol Cybern* 1986; 55: 71–82. doi:10.1007/BF00341922
- [23] Ronneberger O, Fischer P, Brox T et al. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, (eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 Cham: Springer International Publishing*; 2015: 234–241. doi:10.48550/arXiv.1505.04597
- [24] Goodfellow I, Pouget-Abadie J, Mirza M et al. Generative Adversarial Nets. In: Ghahramani Z, Welling M, Cortes C, (eds.) *Curran Associates, Inc.*; 2014. doi:10.48550/arXiv.1406.2661
- [25] Zhu B, Liu JZ, Cauley SF et al. Image reconstruction by domain-transform manifold learning. *Nature* 2018; 555: 487–492. doi:10.1038/nature25988
- [26] Häggström I, Schmidtlein CR, Campanella G et al. DeepPET: A deep encoder-decoder network for directly solving the PET image reconstruction inverse problem. *Med Image Anal* 2019; 54: 253–262. doi:10.1016/j.media.2019.03.013
- [27] Liu Z, Chen H, Liu H et al. Deep Learning Based Framework for Direct Reconstruction of PET Images. In: Shen D, Liu T, Peters TM, (eds.) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019 Cham: Springer International Publishing*; 2019: 48–56. doi:10.1007/978-3-030-32248-9_6
- [28] Whiteley W, Luk WK, Gregor J. DirectPET: full-size neural network PET reconstruction from sinogram data. *J Med Imaging* 2020; 7: 1. doi:10.1117/1.JMI.7.3.032503
- [29] ComtatBataille F, Michel C et al. OSEM-3 D Reconstruction Strategies for the ECAT HRRT. In: *IEEE Symposium Conference Record Nuclear Science 2004 Rome, Italy: IEEE*; 2004: 3492–3496. doi:10.1109/NSSMIC.2004.1466639
- [30] Hu Z, Xue H, Zhang Q et al. DPIR-Net: Direct PET Image Reconstruction Based on the Wasserstein Generative Adversarial Network. *IEEE Trans Radiat Plasma Med Sci* 2021; 5: 35–43. doi:10.1109/TRPMS.2020.2995717
- [31] Ma R, Hu J, Sari H 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: 4464–4477. doi:10.1007/s00259-022-05861-2
- [32] Gong K, Guan J, Kim K et al. Iterative PET Image Reconstruction Using Convolutional Neural Network Representation. *IEEE Trans Med Imaging* 2019; 38: 675–685. doi:10.1109/TMI.2018.2869871
- [33] Ulyanov D, Vedaldi A, Lempitsky V. Deep Image Prior. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)2018*
- [34] Gong K, Catana C, Qi J et al. Direct Reconstruction of Linear Parametric Images From Dynamic PET Using Nonlocal Deep Image Prior. *IEEE Trans Med Imaging* 2022; 41: 680–689. doi:10.1109/TMI.2021.3120913
- [35] Yokota T, Kawai K, Sakata M et al. Dynamic PET Image Reconstruction Using Nonnegative Matrix Factorization Incorporated With Deep Image Prior. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)2019*
- [36] Ote K, Hashimoto F, Onishi Y et al. List-Mode PET Image Reconstruction Using Deep Image Prior. *IEEE Trans Med Imaging* 2023; 42: 1822–1834. doi:10.1109/TMI.2023.3239596
- [37] Fessler JA. Optimization Methods for Magnetic Resonance Image Reconstruction: Key Models and Optimization Algorithms. *IEEE Signal Process Mag* 2020; 37: 33–40. doi:10.1109/MSP.2019.2943645
- [38] Gong K, Catana C, Qi J et al. PET Image Reconstruction Using Deep Image Prior. *IEEE Trans Med Imaging* 2019; 38: 1655–1665. doi:10.1109/TMI.2018.2888491
- [39] Chun Y, Fessler JA. Deep BCD-Net Using Identical Encoding-Decoding CNN Structures for Iterative Image Recovery. In: *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP) Zagorochoria: IEEE*; 2018: 1–5
- [40] Gong K, Wu D, Kim K et al. MAPEM-Net: an unrolled neural network for Fully 3 D PET image reconstruction. In: Matej S, Metzler SD, (eds.) *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine Philadelphia, United States: SPIE*; 2019: 102

- [41] Mehranian A, Reader AJ. Model-Based Deep Learning PET Image Reconstruction Using Forward–Backward Splitting Expectation–Maximization. *IEEE Trans Radiat Plasma Med Sci* 2021; 5: 54–64. doi:10.1109/TRPMS.2020.3004408
- [42] Corda-D'Incan G, Schnabel JA, Hammers A et al. Single-Modality Supervised Joint PET-MR Image Reconstruction. *IEEE Trans Radiat Plasma Med Sci* 2023; 1–1. doi:10.1109/TRPMS.2023.3283786
- [43] Whiteley W, Gregor J. CNN-based PET sinogram repair to mitigate defective block detectors. *Phys Med Biol* 2019; 64: 235017. doi:10.1088/1361-6560/ab4919
- [44] Kaplan S, Zhu YM. Full-Dose PET Image Estimation from Low-Dose PET Image Using Deep Learning: a Pilot Study. *J Digit Imaging* 2019; 32: 773–778. doi:10.1007/s10278-018-0150-3
- [45] Wang Y, Zhou L, Yu B et al. 3 D Auto-Context-Based Locality Adaptive Multi-Modality GANs for PET Synthesis. *IEEE Trans Med Imaging* 2019; 38: 1328–1339. doi:10.1109/TMI.2018.2884053
- [46] Lu W, Onofrey JA, Lu Y et al. An investigation of quantitative accuracy for deep learning based denoising in oncological PET. *Phys Med Biol* 2019; 64: 165019. doi:10.1088/1361-6560/ab3242
- [47] Zhou L, Schaefferkoetter JD, Tham IWK et al. Supervised learning with cyclegan for low-dose FDG PET image denoising. *Med Image Anal* 2020; 65: 101770. doi:10.1016/j.media.2020.101770
- [48] Matej S, Surti S, Jayanthi S et al. Efficient 3-D TOF PET Reconstruction Using View-Grouped Histo-Images: DIRECT—Direct Image Reconstruction for TOF. *IEEE Trans Med Imaging* 2009; 28: 739–751. doi:10.1109/TMI.2008.2012034
- [49] Whiteley W, Panin V, Zhou C et al. FastPET: Near Real-Time Reconstruction of PET Histo-Image Data Using a Neural Network. *IEEE Trans Radiat Plasma Med Sci* 2021; 5: 65–77. doi:10.1109/TRPMS.2020.3028364
- [50] Shiri I, Arabi H, Geramifar P et al. Deep-JASC: joint attenuation and scatter correction in whole-body 18F-FDG PET using a deep residual network. *Eur J Nucl Med Mol Imaging* 2020; 47: 2533–2548. doi:10.1007/s00259-020-04852-5
- [51] Li T, Zhang M, Qi W et al. Motion correction of respiratory-gated PET images using deep learning based image registration framework. *Phys Med Biol* 2020; 65: 155003. doi:10.1088/1361-6560/ab8688
- [52] Ledig C, Theis L, Huszar F et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Honolulu, HI: IEEE; 2017: 105–114
- [53] YangJianchaoWright J, Huang TS et al. Image Super-Resolution Via Sparse Representation. *IEEE Trans Image Process* 2010; 19: 2861–2873. doi:10.1109/TIP.2010.2050625
- [54] Jin W, Li X, Fatehi M et al. Guidelines and evaluation of clinical explainable AI in medical image analysis. *Med Image Anal* 2023; 84: 102684. doi:10.1016/j.media.2022.102684
- [55] Ribeiro MT, Singh S, Guestrin C. „Why Should I Trust You?“. Explaining the Predictions of Any Classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco California USA: ACM; 2016: 1135–1144
- [56] Herington J, McCradden MD, Creel K et al. Ethical Considerations for Artificial Intelligence in Medical Imaging: Deployment and Governance. *J Nucl Med* 2023. doi:10.2967/jnumed.123.266110
- [57] Langlotz CP, Allen B, Erickson BJ et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology* 2019; 291: 781–791. doi:10.1148/radiol.2019190613