REVIEW ARTICLE

Explainable Artificial Intelligence in Nuclear Medicine: Advancing Transparency in PET and SPECT Imaging and Radiation Therapy

Hossein Arabi [1*], Masoud Noroozi [2], Hamed Aghapanah [3], Sayna Jamaati [4], Ali Saeeidi Rad [2], Soroush Salari [5], Jafar Majidpour [6], Sirwan Maroufpour [7], Habibollah Dadgar [8], Francesca Russo [9], Andrea Cimini [9]

[1] *Division of Nuclear Medicine & Molecular Imaging, Geneva University Hospital, CH-1211, Geneva, Switzerland*

[2] *Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran*

[3] *School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran*

[4] *Department of Energy Engineering, Sharif University of Technology, Tehran, Iran*

[5] *Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran*

[6] *Department of Software Engineering, College of Engineering, University of Raparin, Ranya, Kurdistan Region, Iraq*

[7] *Medical Physics Group, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran*

[8] *Cancer Research Center, RAZAVI Hospital, Imam Reza International University, Mashhad, Iran*

[9] *Nuclear Medicine Unit, St. Salvatore Hospital, 67100 L'Aquila, Italy*

*Corresponding Author: Hossein Arabi

Email: hossein.arabi@unige.ch

Received: 15 November 2025 / Accepted: 06 December 2025

Abstract

The integration of Artificial Intelligence (AI) into nuclear medicine has transformed diagnostic and therapeutic processes, yet the opaque nature of many AI models hinders clinical adoption and trust. This narrative review aims to synthesize the current landscape of explainable AI (XAI) in nuclear medicine, emphasizing its role in enhancing transparency, bias mitigation, and regulatory compliance for robust clinical integration. Key chapters cover the fundamentals of XAI in nuclear medicine; XAI applications in PET and SPECT instrumentation and acquisition; image reconstruction; quantitative imaging and corrections; post-reconstruction processing and analysis; and radiotherapy. The review concludes with a discussion of challenges, limitations, and future directions, advocating for interdisciplinary advancements to bridge AI innovation with practical utility in patient care.

Keywords: Explainable Artificial Intelligence; Artificial Intelligence; Positron Emission Tomography; Single Photon Emission Computed Tomography; Nuclear Medicine.

DOI: https://doi.org/10.18502/fbt.v13i1.20789

1. Introduction

The integration of Artificial Intelligence (AI) into nuclear medicine has revolutionized diagnostic and therapeutic workflows, particularly in Positron Emission Tomography (PET) and single-photon emission computed tomography (SPECT) imaging, as well as radionuclide therapy [1, 2]. While foundational AI concepts trace back to the 1950s, the surge in Deep Learning (DL) methodologies since the 2010s has propelled applications in multimodal imaging, enabling automated processes that surpass traditional human-dependent tasks such as lesion detection, image reconstruction, and quantitative analysis [2]. In nuclear medicine, DL excels at managing high-dimensional datasets from hybrid systems (e.g., PET/CT or SPECT/MR), mitigating noise, optimizing radiation doses, and enhancing temporal resolution in dynamic studies, all while adapting to patient-specific variability and reducing inter-observer bias [3, 4].

Despite these advancements, the opaque "black-box" nature of many DL models, wherein decision pathways remain opaque, poses critical barriers in clinical settings, where transparency is paramount for regulatory compliance (e.g., FDA guidelines), clinician adoption, and ethical patient care [5]. Explainable AI (XAI) emerges as a pivotal solution, rendering model outputs interpretable through techniques such as feature attribution (e.g., SHapley Additive exPlanations [SHAP] or Local Interpretable Model-agnostic Explanations [LIME]), uncertainty quantification (e.g., Bayesian approaches or Monte Carlo dropout), saliency maps, and physics-informed architectures that incorporate domain-specific priors like tracer kinetics or photon transport models [6, 7]. In PET and SPECT, XAI facilitates trustworthy applications across the imaging pipeline: from instrumentation enhancements (e.g., timing calibration and motion correction) to reconstruction algorithms that provide pixel-wise uncertainty maps, quantitative corrections for attenuation and scatter, post-reconstruction segmentation with visual attributions, and dosimetry in radionuclide therapy (e.g., voxel-level dose predictions for [^{177}Lu]-based treatments) [8, 9]. For instance, XAI can attribute predictions to physiologically relevant features, such as metabolic hotspots in oncology or kinetic parameters in neurology, thereby fostering clinician-model collaboration and improving outcomes in theranostics [10-12].

A standard nuclear medicine workflow encompasses examination planning, acquisition, interpretation, and reporting, with AI augmenting each phase, either by automating routine tasks to alleviate workload or by achieving superhuman precision in detecting subtle pathologies like micrometastases or predicting therapeutic responses [13]. Recent innovations, including inter-modality translation (e.g., MRI-to-CT for attenuation maps) and low-dose enhancements, underscore AI's transformative potential, yet highlight the imperative for XAI to ensure reproducibility, bias mitigation, and alignment with physical principles [14-20].

This narrative review synthesizes the evolving landscape of XAI in nuclear medicine, with a primary emphasis on PET/SPECT imaging and radionuclide therapy. Drawing from recent literature, we delineate foundational principles, key applications, inherent challenges, and prospective directions to guide interdisciplinary efforts toward robust, transparent AI integration in clinical practice.

2. Fundamentals of Explainable AI in Nuclear Medicine

In nuclear medicine, where high-stakes decisions rely on interpreting complex, noisy datasets from PET, SPECT, and radiation therapy, the "black-box" opacity of traditional DL models poses significant hurdles [5, 21, 22]. Opaque AI systems, such as Convolutional Neural Networks (CNNs) trained on volumetric PET/SPECT images, excel at tasks like lesion detection, image reconstruction, and dosimetry but obscure internal decision pathways, leading to challenges in clinical validation, regulatory approval (e.g., under FDA's AI/ML-enabled medical device guidelines), and bias detection [5, 21, 23, 24]. For instance, DL models may inadvertently amplify artifacts from photon noise, scatter, or attenuation in low-count PET acquisitions, resulting in non-reproducible predictions without traceable reasoning [25-30]. In contrast, XAI frameworks demystify these processes by providing interpretable outputs, such as feature attributions or uncertainty maps, that align with domain-specific physics (e.g., tracer kinetics or

Monte Carlo-based photon transport simulations) [5, 21].

XAI methods are broadly categorized into post-hoc (applied after model training, e.g., model-agnostic perturbations) and ad-hoc (intrinsically built-in, e.g., attention layers), with adaptations for nuclear medicine's unique data characteristics: high-dimensional 3D/4D volumes, inherent Poisson noise, and multimodal fusion (e.g., PET/CT or SPECT/MR) [5, 21, 23]. These enable clinicians to verify model focus on physiologically relevant features, like metabolic hotspots in oncology or perfusion deficits in cardiology, fostering trust and integration into workflows [5, 21, 31].

Core Principles of XAI

XAI principles derive from cognitive science and game theory, distinguishing interpretability (inherently transparent models, e.g., decision trees) from explainability (post-hoc rationales for black-boxes). Core taxonomy classifies by: (A) Scope: local (per-instance) vs. global (model-level); (B) Timing: ante-hoc (pre-training), intrinsic/ad-hoc (design-embedded), post-hoc (after); (C) Granularity: signal-level (voxels) to concept-level (e.g., "hotspot").

Key desiderata (properties for trustworthy XAI): (1) Fidelity (matches model); (2) Stability (consistent inputs); (3) Plausibility (human-aligned); (4) Robustness (adversarial-resistant); (5) Simulatability (mental replay). In nuclear medicine, prioritize stability to combat scan noise (e.g., via robustness to Gaussian perturbations) and plausibility to ensure physics alignment (e.g., explanations mirroring signal decay models), fostering clinician trust. Evaluation blends quantitative (AOPC (Area Over Perturbation Curve) for sufficiency via ranked voxel perturbations; Infidelity for faithfulness via physics-like noise) and qualitative (clinician surveys for plausibility) (Table 1) [32-34].

Post-hoc XAI techniques, dominant in nuclear medicine due to their flexibility with pre-trained DL models, generate explanations by analyzing model outputs retrospectively, often through gradient propagation or perturbations tailored to handle nuclear medicine's noisy, volumetric data [5, 21, 23, 35, 36]. Gradient-based methods, such as saliency maps and Gradient-weighted Class Activation Mapping (Grad-CAM), compute voxel-wise importance by back-propagating gradients from the output layer to highlight influential regions in PET/SPECT images [23, 37, 38]. For example, Gradient-based methods like Grad-CAM improve interpretability in noisy environments, such as myocardial perfusion SPECT where low-dose acquisitions can introduce streak artifacts due to reduced photon counts; variants like Grad-CAM++ further enhance spatial resolution by incorporating pixel-wise weighting for multiple object instances, potentially suppressing non-relevant artifacts [23, 38-40]. Layer-wise Relevance Propagation (LRP) and Deep Learning Important FeaTures (DeepLIFT) extend this by conserving relevance scores across layers, addressing gradient saturation in deep networks, LRP, for instance, has been adapted for dopamine transporter SPECT in Parkinson's disease classification, visualizing striatal uptake patterns in 3D volumes by propagating relevance voxel-by-voxel, revealing model reliance on binding ratios amid noise [41-43]. Perturbation-based approaches like Local Interpretable Model-agnostic Explanations (LIME) approximate local decision boundaries by fitting interpretable surrogates (e.g., linear models) to perturbed super-voxels, grouping similar pixels to manage nuclear medicine's high dimensionality; LIME's efficiency shines in

Table 1. Classification of XAI Methods

Scope	Timing	Method Category	Core Principle	Nuclear medicine Adaptation/Evaluation
Local/Global	Post-hoc	Gradient (Grad-CAM)	Backprop sensitivities	Voxel heatmaps; AOPC
Local	Post-hoc	Perturbation (LIME)	Surrogate fitting	Super-voxels; Stability
Local/Global	Post-hoc	Game-theoretic (SHAP)	Marginal contributions	Radiomics; Infidelity
Local/Global	Intrinsic	Attention	Dynamic weighting	3D kinetics; Plausibility
Global	Intrinsic	PINNs	Physics-constrained loss	Uncertainty maps; Fidelity

volumetric imaging, where random perturbations quantify contributions from noisy regions, as demonstrated in CT for COVID-19 classification via occlusion-based explanations of lung opacities, and in SPECT for neurological disease detection, with analogous potential for extension to PET tumor segmentation by perturbing metabolic features to assess boundary impacts [5, 35, 44, 45].

SHapley Additive exPlanations (SHAP), a game-theoretic post-hoc method, provides robust, additive feature attributions by computing average marginal contributions across coalitions, offering both local (instance-specific) and global (model-wide) insights particularly suited for nuclear medicine's correlated radiomic features [46]. In PET/CT radiomics for lung cancer nodal staging, SHAP ranks features like SUVmax (from PET) and texture metrics (e.g., Gray Level Co-occurrence Matrix correlation from CT), revealing contributions to metastasis prediction due to tumor heterogeneity; this disentangles multimodal synergies in noisy fusions, with SHAP outperforming LIME in stability for high-dimensional data. DeepSHAP, an approximation for deep learning, has been applied to mapping multivariate relationships in amyloid PET for Alzheimer's cognition prediction, attributing regions to deposition patterns while handling variability via reference baselines (e.g., healthy scans), though sensitive to anatomical differences [47, 48].

Ad-hoc XAI methods embed interpretability during model design, yielding intrinsically transparent architectures better resilient to NM's challenges, such as volumetric deformations and low signal-to-noise ratios in SPECT. Attention mechanisms, integrated via Transformer layers or self-attention modules, dynamically weight relevant features, producing attention maps that visualize information flow, in SPECT for Parkinson's, attention-based DenseNet models focus on basal ganglia regions (including nigrostriatal pathways), suppressing noise from surrounding tissues by computing softmax-normalized scores across patches, adaptable to 3D volumes through multi-head attention for capturing long-range dependencies in DaTSCAN tracer uptake. Capsule networks, which encode hierarchical entity representations via vector outputs and dynamic routing, provide equivariant explanations robust to rotations in volumetric data; though less common in NM, they mirror adaptations from CT nodule detection, where capsules disentangle pose-invariant features in noisy lung images, with potential extensions to SPECT for dosimetry planning. Physics-informed neural networks (PINNs) further tailor ad-hoc XAI by constraining losses with nuclear medicine priors (e.g., radiative transfer equations analogous to Boltzmann transport for scatter correction), generating uncertainty-quantified outputs like Bayesian dropout maps for voxel-level dose predictions in radiotherapy, contrasting opaque models by enforcing physical plausibility. For example, attention-augmented models in multimodal PET/MRI for Alzheimer's localize atrophy-metabolism mismatches, with rollout visualizations tracing layer-wise contributions amid noise [49-51].

Figure 1 highlights the fundamental explainability-performance trade-off in AI systems: as models prioritize superior accuracy (e.g., via complex deep learning or ensemble techniques), their decision-making processes become increasingly opaque ("black-box"). Conversely, transparent models like decision trees or SVMs sacrifice some performance for inherent interpretability. In medical imaging contexts, such as nuclear medicine diagnostics, this tension underscores the value of XAI strategies (e.g., post-hoc attributions like Grad-CAM or SHAP) to "unlock" high-performing deep models, enabling clinicians to trust and validate predictions without compromising efficacy.

Despite these foundational principles, XAI in nuclear medicine confronts domain-specific challenges: scalability for 4D dynamic studies (e.g., SHAP coalitions scale, demanding approximations like FastSHAP); fidelity in low-count regimes (Poisson noise erodes attribution stability); and bias in multimodality (e.g., PET/CT misalignments propagate via uncalibrated gradients). Future directions pivot to hybrid physics-XAI frameworks: physics-informed SHAP (PINNs constraining losses with Boltzmann transport, yielding uncertainty-aware attributions for [^{177}Lu]-dosimetry); perturbation hybrids (masking voxels akin to ADNI multimodal drops for global fidelity benchmarks); and real-time intrinsics (quantized attentions for edge-deployed SPECT, aligning with FDA's AI/ML SaMD guidelines). These

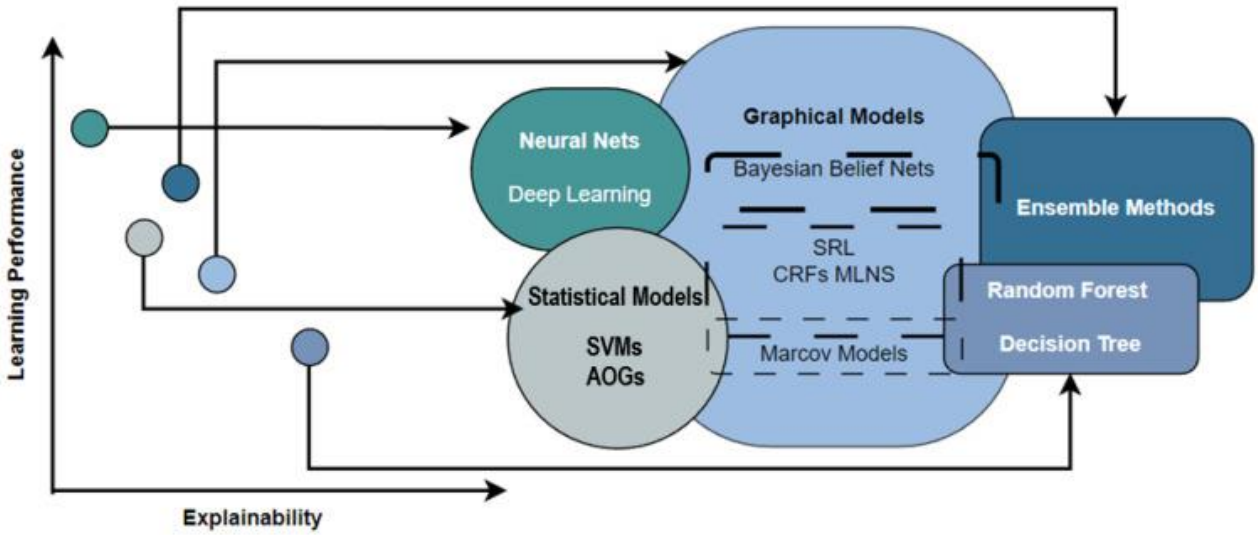

Figure 1. Trade-off curve illustrating the inverse relationship between model explainability (x-axis) and learning performance (y-axis) in machine learning paradigms. Highly interpretable models, such as simpler statistical approaches, cluster toward greater explainability but lower predictive accuracy. In contrast, high-performance architectures like deep neural networks and ensemble methods exhibit reduced transparency. Key categories include: Statistical Models [SVMs: Support Vector Machines; AOGs: And-Or Graphs]; Graphical Models [Bayesian Belief Nets; SRL: Statistical Relational Learning; CRFs: Conditional Random Fields; MLNs: Markov Logic Networks; Markov Models]; Neural Nets [encompassing Deep Learning]; and Ensemble Methods [Random Forest; Decision Tree]. Reprinted from [52] under CC BY 4.0 copyright

promise trustworthy, physics-plausible nuclear medicine AI, bridging principles to practice [5].

3. Explainable AI in PET and SPECT Instrumentation and Acquisition

In PET and SPECT systems, the foundational processes of photon detection, signal processing, and data acquisition directly dictate image quality, quantitative accuracy, and clinical utility. These modalities rely on scintillation crystals (e.g., LYSO:Ce for PET, NaI(Tl) for SPECT) coupled to photosensors like silicon photomultipliers (SiPMs), where raw waveforms encode critical information on energy, position, and timing amid high noise, pile-up events, and physical imperfections such as Depth-Of-Interaction (DOI) variations or collimator septal penetration. Conventional AI, particularly DL models like CNNs, has revolutionized low-level tasks, yielding 20-40% gains in timing precision, positioning accuracy, and calibration stability, but often operates as "black boxes," obscuring decision rationales in noisy, physics-constrained environments. This opacity risks propagating untraceable errors, especially in low-light yields (~10-30 photons/keV) or high-activity scenarios, complicating validation against physical principles and regulatory scrutiny (e.g., FDA's AI/ML-based SaMD framework (proposed regulatory framework for modifications to artificial intelligence/machine learning-based software as a medical device)). XAI provides transparency and enables physicists to dissect AI contributions, confirm physics compliance, and iteratively optimize hardware/acquisition.

The integration of XAI into PET and SPECT instrumentation represents an emerging frontier in nuclear medicine, where transparency is paramount for validating AI decisions against fundamental physical principles. Unlike the mature application of XAI in downstream tasks such as lesion detection, image reconstruction, and diagnostic classification (covered in the following sections), where methods like SHAP and Grad-CAM elucidate lesion attributions in reconstructed images, low-level instrumentation processes (e.g., raw photon detection, signal processing, and hardware calibration) remain underexplored. Here, AI excels in handling noisy scintillator waveforms from silicon photomultipliers (SiPMs), achieving 20-40% improvements in timing resolution, event positioning, and depth-of-interaction

(DOI) decoding, but black-box models risk untraceable errors in low-photon-yield (~10-30 photons/event) or high-pileup environments. XAI bridges this by dissecting feature contributions (e.g., waveform rise time vs. decay tail), ensuring physics compliance, and enabling iterative hardware design. Seminal works are scarce, focusing primarily on PET time-of-flight (TOF) estimation, while XAI applications are accompanied by recommendations for wider use [53-60].

Pioneering efforts demonstrate XAI's transformative potential in TOF-PET timing. Naunheim *et al.* [54] introduced a gradient-boosted decision tree (GBDT) model augmented with residual physics constraints, modeling nonlinear timewalk as $\Delta t\_phys = f(energy, position)$, to process raw LYSO:SiPM waveforms (Figure 2). The objective was to surpass conventional constant-fraction discriminators amid Poisson noise and variable photon statistics, targeting sub-200 ps Coincidence Time Resolution (CTR) for clinical total-body PET. Achieved a 21% CTR gain (235 ps to 185 ps) in 19-mm slab detectors using list-mode data. XAI via SHAP provided global/local attributions: 81% importance to measured Δt and photon count, revealing learned higher-order corrections without opacity; tree visualizations traced paths to physical priors. This intrinsic/post-hoc hybrid validated model fidelity against Monte Carlo simulations (GATE), paving the way for scanner-agnostic deployment.

Complementing this, Petersen *et al.* [61] developed a DL framework for event positioning and inter-crystal scatter rejection in light-sharing, depth-encoding PET detectors (e.g., Prism-PET modules). The objective was to mitigate parallax errors (~5-10 mm) and scatter-induced blurring by decoding DOI and interaction coordinates from multiplexed SiPM signals, improving spatial resolution by 15-25% over Anger logic. While DL excelled, the authors explicitly highlighted XAI's future necessity for clinician trust, noting post-hoc tools (e.g., integrated gradients) could map decisions to scatter paths (Klein-Nishina compliant). No XAI was implemented, underscoring the niche's nascency.

Despite these pioneering efforts, significant gaps in XAI application to PET and SPECT instrumentation highlight ripe opportunities for innovation. In SPECT, where collimator septal penetration and Compton scatter distort ~10-20% of events, no XAI frameworks yet interpret gamma event localization in ML models, DenseNets could leverage Grad-CAM heatmaps to visualize penetration paths aligned with Klein-Nishina physics, guiding collimator refinements. Similarly, deep learning excels in DOI decoding for light-sharing/monolithic setups, yet lacks XAI to unpack multi-layer waveform attributions amid variable light

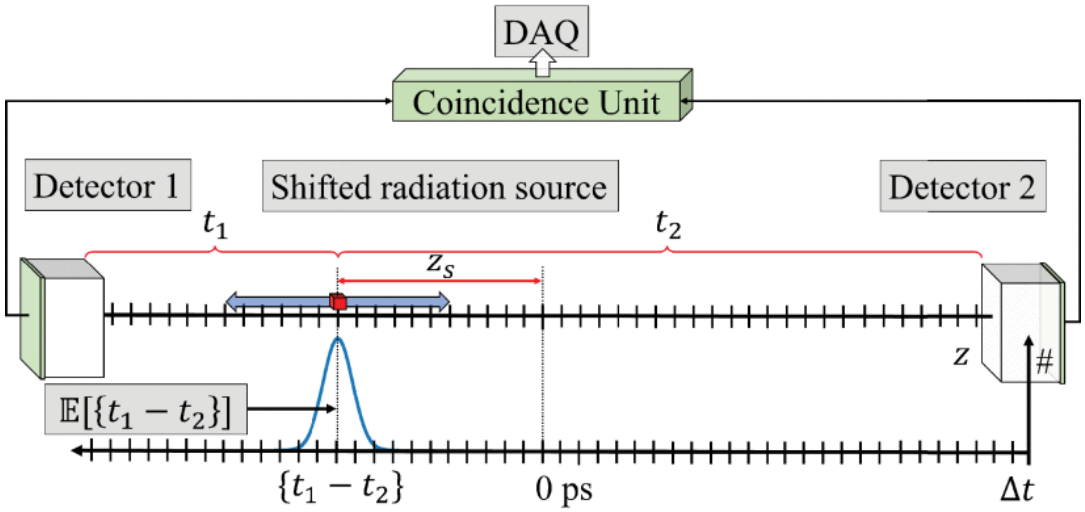

Figure 2. This schematic depicts the labeling procedure for generating annotated datasets in supervised learning. The radiation source (red cube) is methodically repositioned at discrete locations $ z_s $ along the central z-axis of the coordinate system. These shifts produce distinct γ-photon travel times $ t_1 $ and $ t_2 $ to paired detectors. The expected time difference $ \mathbb{E[t_1 - t_2] $ is computed as the precise ground-truth label for each configuration. This physics-driven method yields scalable, noise-aware labels, empowering models to infer source positions from real-time time-of-flight data. Reprinted from [54] under CC BY 4.0 copyright

yields. Event positioning via CNNs achieves sub-mm precision in monolithic crystals, but SHAP could dissect edge blurring from scatter, while SiPM gain drifts await interpretable recalibration tracing thermal/aging effects for scanner longevity. Hardware optimization via physics-informed NNs (PINNs) models can lead to promising development in PET and SPECT instrumentation.

As total-body PET and SPECT systems continue to push the boundaries of sensitivity (e.g., >100× gains in human explorers like PennPET and EXPLORER) and spatial/temporal resolution (sub-200 ps CTR, mm-scale DOI), embedding XAI directly into instrumentation, via lightweight, real-time methods like SHAP-on-FPGA or intrinsic GBDTs, will cultivate inherently trustworthy AI by demystifying decisions at the photon level, from waveform parsing to adaptive acquisition. This transparency empowers physicists and engineers with actionable insights, such as feature heatmaps linking SiPM drifts to CTR degradation or Grad-CAM overlays revealing septal penetration artifacts, thereby fueling rapid iterative design cycles [62, 63].

4. Explainable AI in Image Reconstruction

PET and SPECT image reconstruction has traditionally relied on analytical methods like filtered backprojection (FBP) and iterative algorithms such as ordered subset expectation maximization (OSEM). While effective, these approaches suffer from limitations including noise amplification, reduced contrast recovery, and sensitivity to artifacts from attenuation, scatter, or motion. AI, particularly DL, has revolutionized this field by framing reconstruction as an image-to-image translation task [64-68]. DL models, including CNNs, GANs, and unrolled iterative networks, enable superior resolution, contrast enhancement, noise suppression, and artifact correction, often outperforming conventional methods by 10-20% in quantitative metrics like normalized root-mean-square error (NRMSE) [69].

Hybrid and physics-informed approaches dominate recent advances: unrolled networks (e.g., unfolding OSEM/MAP-EM into recurrent layers) embed forward models for data fidelity, while post-hoc DL (e.g., U-Net denoising) refines outputs. In PET, kernel priors support list-mode, motion-corrected, and parametric imaging [70-72]. Clinically approved tools now integrate multimodal data, promising routine deployment. However, challenges like training data scarcity and hallucination risks remain.

The Imperative for Explainable AI (XAI) in Reconstruction

XAI addresses the "black-box" critique of DL recon by providing interpretable insights into model decisions, fostering clinician trust and regulatory approval. In PET/SPECT, where artifacts mimic lesions or obscure uptake, XAI overlays (e.g., attention maps, SHAP values) elucidate how networks reduce noise, correct misalignments, or recover lesions, critical for reducing false positives/negatives and improving detectability. User-centric XAI aligns with workflows, enabling rapid validation and hybrid human-AI decisions.

In this regard, pioneering physics-informed unrolled networks exemplify interpretable recon. Mehranian and Reader's FBSEM-Net unrolls forward-backward splitting expectation-maximization, replacing priors with learnable CNN residuals while fixing physics operators (e.g., system matrix), inherently traceable via iteration blocks mirroring OSEM [73]. This hybrid outperforms OSEM/MAPEM (NRMSE ~14% vs. 21%) in low-dose PET/MR brain data. Corda-D'Incan *et al.* [74] extend it with iteration-dependent targets/losses and sequential training, slashing memory by 98% for 3D fully unrolled nets (up to 100+ iterations), boosting generalization without leapfrogging artifacts.

Physics-informed NNs (PINNs) advanced dynamic PET as Ferrante *et al.* [50] embed AIF shape priors in 3D depth-wise CNNs for metabolite-corrected plasma input (Figure 3), yielding Pearson r=0.89 vs. invasive sampling, explainable via PDE residuals. Salomonsen *et al.* [75] apply PINN-CycleGAN for voxel-wise kinetics, predicting AIFs with parameter maps rivaling references, interpretable through cycle-consistency losses enforcing biophysics.

Direct XAI applications target recon artifacts/interpretation: Champendal *et al.* [76] define user-centric XAI criteria for PET/CT denoising, dual-level explanations (global confidence scores + case-specific "what/when/how" visuals) preserve

workflows. Salimi *et al.* [77] use organ segmentation (lungs/liver) + random forest for respiratory misalignment detection (AUC=0.93), mimicking radiologist contour checks. Miller *et al.* [78] demonstrate explainable DL (heatmaps) boosting physician MPI accuracy (AUC 0.747 to 0.779) in SPECT/PET perfusion. Reviews like Apostolopoulos *et al.* [79] synthesize DL-SPECT cardiac recon, urging XAI (e.g., Grad-CAM) for denoising/attenuation.

Future Directions

XAI in PET/SPECT recon remains niche, with <5% of DL papers incorporating it, demanding dedicated modeling like SHAP/Grad-CAM on unrolled layers, adversarial XAI for hallucinations, and prospective trials. Hybrid OSEM+DL+XAI pipelines, multimodal fusion (e.g., PET/CT/MR), and real-time clinician dashboards will drive adoption, reducing artifacts while enhancing lesion detection. Collaborative benchmarks and FDA-guided validation are essential to transition from research to routine.

5. Explainable AI in Quantitative Imaging and Corrections

Attenuation and Scatter Correction

Attenuation correction (AC) and scatter correction (SC) are cornerstone processes for achieving quantitative accuracy in PET and SPECT imaging, enabling reliable standardized uptake value (SUV) measurements critical for oncology tumor staging and neurology neurodegenerative assessments. While DL has proliferated for these tasks, exemplified by indirect μ-map generation from non-attenuated corrected (NAC) emission data or direct AC image prediction, the majority remain opaque black-box models [28, 80-82]. This opacity hampers clinical adoption, as clinicians cannot discern correction artifacts (e.g., overestimation in inferior walls from uncorrected diaphragmatic attenuation) or trust quantitative outputs in heterogeneous oncology lesions. XAI addresses this by providing interpretable heatmaps or relevance scores, yet XAI integration in PET/SPECT AC/SC remains niche, with only a few DL works incorporating techniques like Grad-CAM or SHAP.

Pioneering XAI applications have emerged in synthetic CT (sCT) generation from MR images for PET/MR AC, where attention mechanisms guide bone/soft-tissue delineation to minimize SUV bias (<3%). Dovletov *et al.* [83] proposed a Grad-CAM-guided U-Net, using class-specific activation maps to focus translation on bone/air regions, yielding 2-5 dB PSNR gains and 10% MAE reduction versus baseline

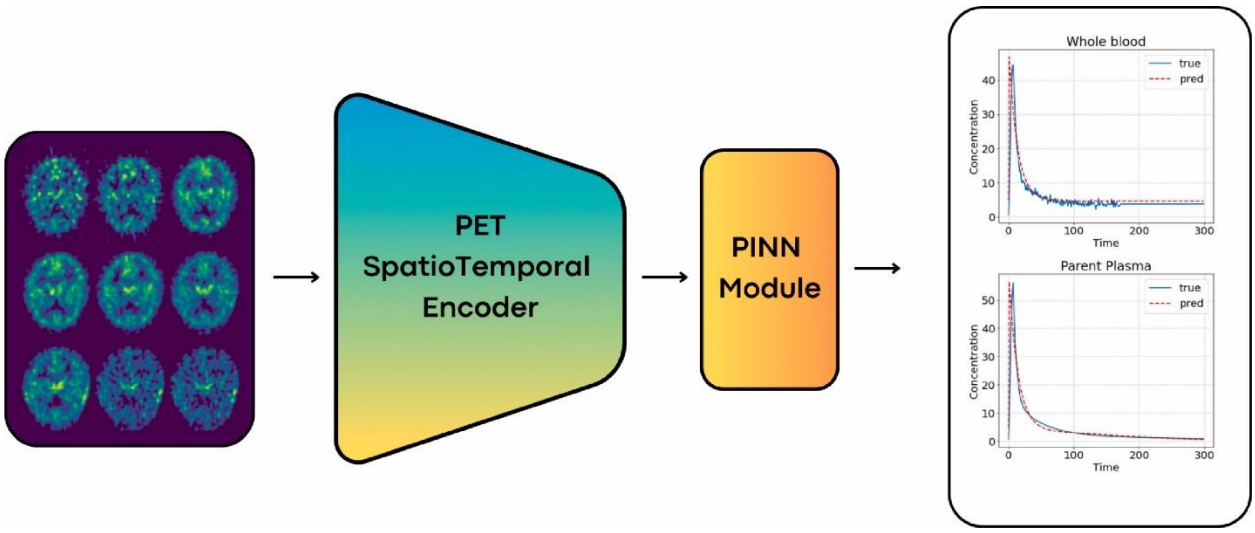

Figure 3. Overview of the deep learning architecture for arterial input function (AIF) estimation in dynamic PET imaging. This schematic provides an illustrative overview of the proposed architecture's data flow, starting from the input layers and advancing through successive computational stages to the final estimation of whole blood and parent plasma input curves. Each block denotes a distinct computational unit or layer, engineered to capture key spatio-temporal features from 4D PET datasets. These features are then channeled into the Physically-Informed Neural Network (PINN) module, which enforces domain-specific physical constraints alongside curve morphology priors. Reprinted from [50] under CC BY 4.0 copyright

U-Net on public RIRE datasets, ideal for radiation-free AC in brain oncology. Building thereon, Dovletov *et al.* [84] introduced double Grad-CAM (coarse + fine-grained) for enhanced sCT in pelvis/head, improving SSIM by 0.05 and bone accuracy to 99%, preventing harmful dosimetry errors. Complementary works include Chen *et al.*'s [85] SHAP-like visualizations for cross-vendor SPECT μ-map synthesis from emissions (NMSE 5.1%) and Shi *et al.*'s [86] cGAN with interpretable outputs (nMAE 3.6%), extending XAI to CT-free SPECT.

Low-Dose and Fast-Acquisition Imaging

AI, particularly DL, has revolutionized image reconstruction in low-dose PET and SPECT imaging, enabling significant reductions in radiation exposure and acquisition times while preserving diagnostic quality [25, 87, 88]. In low-dose PET, CNNs, U-Net architectures, GANs, and more recently diffusion models have been extensively employed to denoise and reconstruct standard-dose-like images from low-count sinograms or projections, outperforming traditional iterative methods like OSEM [89, 90]. Similarly, for fast-acquisition or sparse-view SPECT, where fewer projections reduce scan duration and patient motion artifacts, DL-based methods, including GANs and transformer variants, facilitate high-fidelity reconstructions from undersampled data [91, 92].

Despite these advances, XAI remains underexplored in low-dose/fast PET and SPECT reconstruction. While physics-informed neural networks (PINNs) and uncertainty quantification (UQ) via Bayesian methods are emerging, dedicated XAI techniques (e.g., saliency maps, confidence scores) are scarce, with a few publications addressing interpretability in this niche.

In this regard, Vlašić *et al.* [93] pioneered UQ in low-dose PET via deep posterior sampling, a cornerstone of XAI for trustworthiness. Their conditional GAN generator, built on residual-in-residual dense blocks (RRDBs) with StyleGAN-inspired per-pixel noise injection, conditions on MLEM-reconstructed L-PET (or very-low-dose vL-PET) and T1-MRI to sample diverse standard-dose PET realizations. A physics-informed consistency loss (Radon transform alignment) ensures measurement fidelity, alongside diversity and first-moment penalties to prevent mode collapse. On simulated BrainWeb (L/vL-PET) and real ADNI data, it surpasses MLEM and suDNN in PSNR/SSIM (e.g., 31.97 dB/0.9216 ADNI vL-PET) while yielding variance-based uncertainty maps that scale meaningfully with dose, higher in low-count regions for intuitive certainty visualization.

Complementing this, a physics-informed approach proposed by Tang *et al.* [94], which introduces a collaborative paradigm for joint low-dose PET/CT reconstruction. From non-attenuation-corrected low-dose PET (NAC-LPET) and low-dose CT (LCT), their

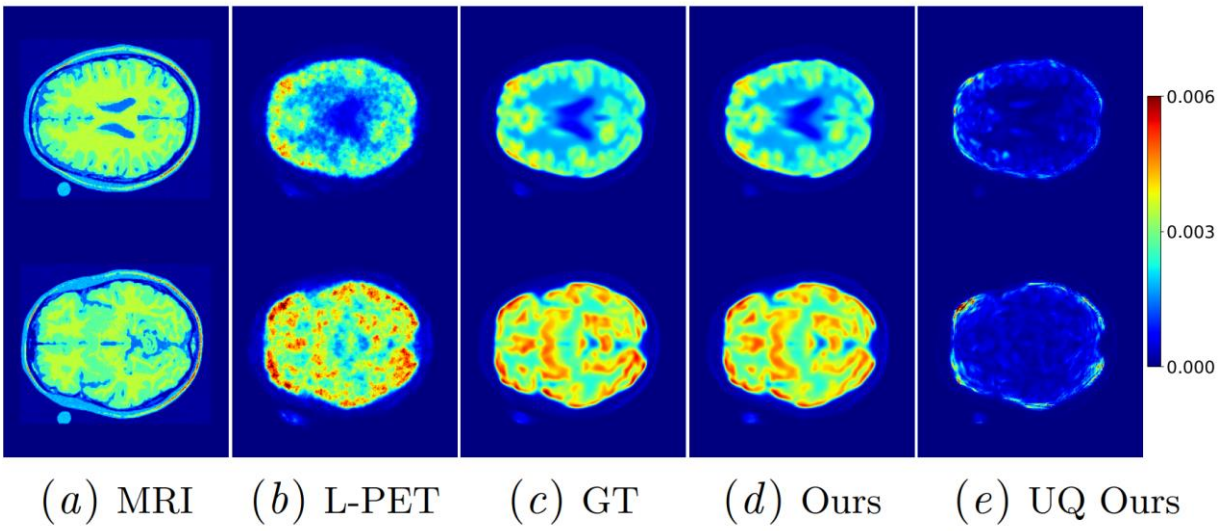

Figure 4. Standard-dose PET (S-PET) reconstructions alongside physically meaningful pixel-wise uncertainty quantification (UQ) maps, generated from low-dose PET (L-PET) and MRI inputs using deep posterior sampling (reproduced from Vlašić *et al.*, [93] under CC-BY 4.0)

coarse-to-fine network reconstructs attenuation-corrected standard-dose PET (AC-SD-PET) and standard-dose CT (SCT). Key innovations include modality-specific Mamba-powered Expert Networks (hybrid U-shaped with tri-oriented Mamba layers for long-range 3D dependencies in whole-body scans) fused via Domain Adapters, and a novel Physics-informed Mutual Loss enforcing PET-CT domain consistency through mutual information minimization. Evaluated on 251 TCIA NSCLC whole-body cases, it achieves state-of-the-art PSNR (36.44 dB PET/37.19 dB CT) and SSIM (96.73%/97.32%), outperforming 3D U-Net, cGAN, DDPM, AIGAN, and MVAE, with ablations confirming each component's value. This enhances explainability by embedding physical priors transparently.

In summary, XAI in low-dose/fast PET/SPECT remains a niche frontier, demanding advanced algorithm development to integrate hybrid physics-DL models, scalable UQ, and clinician-centric visualizations for widespread adoption.

Artifact Reduction and Partial Volume Correction

AI has significantly advanced partial volume correction (PVC) in PET and SPECT imaging, addressing the limitations of traditional methods that often rely on anatomical priors or simplified models. Deep learning approaches have been employed to jointly perform denoising and PVC, enabling more accurate quantification of tracer uptake in small structures like tumors or brain regions [95-97]. For instance, attention-based neural networks have been developed to predict PVC-corrected images directly from non-corrected PET data, bypassing the need for co-registered anatomical images and improving recovery coefficients in low-dose scenarios. Similarly, iterative deep learning frameworks incorporate spatial resolution modeling to restore activity in partial volume-affected areas, demonstrating superior performance over classical techniques like the Yang method in clinical datasets [98].

In artifact correction, AI algorithms excel at detecting and mitigating issues such as halo artifacts, metal-induced distortions, and truncation errors in hybrid PET/CT or PET/MRI scans [65, 99]. Supervised learning models, trained on large datasets of artifact-contaminated images, can automatically identify and correct these anomalies in real-time, enhancing image quality and diagnostic reliability.

For post-reconstruction corrections, particularly motion correction, AI has introduced robust solutions to compensate for patient movement, which can severely degrade image resolution and quantification in dynamic PET/SPECT studies [100]. Deep learning-based methods utilize fast reconstructions as inputs to estimate and correct head or respiratory motion, achieving sub-millimeter accuracy in brain imaging without external hardware [101]. In whole-body applications, unified frameworks integrate motion estimation with reconstruction pipelines, leveraging optical surface information or k-space data to phase-sort and correct respiratory artifacts in SPECT, thereby improving lesion detectability in pulmonary regions [102]. Clinical evaluations of these techniques, such as unsupervised respiratory motion correction (uRMC), have demonstrated enhanced image sharpness and reduced blurring in real-world settings, with potential extensions to cardiac and oncologic imaging [103]. Overall, AI facilitates end-to-end post-processing workflows that combine motion correction with attenuation and scatter adjustments, significantly outperforming traditional gated or rigid registration methods in terms of speed and efficacy [100, 104].

Despite these advancements in AI for PVC, artifact correction, and motion correction in PET/SPECT, the integration of XAI remains underexplored, presenting key research opportunities. Incorporating XAI could provide certainty maps or feature attribution visualizations to elucidate model decisions, fostering clinical trust and enabling validation in high-stakes diagnostic scenarios where transparency is crucial for regulatory approval and error mitigation.

6. Explainable AI in Post-Reconstruction Processing and Analysis

Image Segmentation

Despite the extensive application of DL methods in PET and SPECT segmentation tasks, such as tumor delineation and organ contouring, research integrating XAI remains remarkably limited. DL has revolutionized nuclear medicine imaging by enabling automated, high-precision segmentation that

outperforms traditional threshold-based or atlas-based approaches, particularly in handling noisy, low-resolution data inherent to these modalities [105-111]. Seminal works in DL for PET/SPECT segmentation include the adaptation of U-Net architectures for tumor segmentation in PET/CT scans, wherein these studies highlight DL's efficacy in tasks like lesion detection and quantitative analysis but often treat models as black boxes, underscoring the need for XAI to enhance clinical trust and error debugging.

Three key review articles provide broader context on XAI in medical imaging, with varying relevance to PET and SPECT imaging : Champendal *et al.* [112] conducted a scoping review mapping XAI methods across modalities like MRI, CT, and radiography, identifying visual (e.g., saliency maps) and numerical outputs as dominant for tasks including segmentation, while noting terminology inconsistencies between "explainable" and "interpretable"; Salih *et al.* [113] focused on XAI for cardiac imaging, discussing techniques like Grad-CAM and SHAP for interpreting DL models in CMR and echocardiography segmentation, emphasizing the trade-off between performance and interpretability; and Usmani *et al.* [114] reviewed DL-based segmentation of ^{68}Ga-PSMA PET for prostate cancer tumor volume assessment, highlighting AI's role in radiotherapy planning but touching on XAI via visualization methods like Grad-CAM to explain model focus on metastatic regions, though primarily emphasizing automation over full explainability. One notable original work incorporating XAI in segmentation, albeit for mammograms rather than PET/SPECT, is Farrag *et al.* [115] which proposed a double-dilated CNN to preserve local resolution during tumor segmentation, addressing kernel sparsity issues in dilated convolutions; the model achieved high Dice similarity (0.92) and low miss detection rates, with explainability enhanced through Grad-CAM visualizations to interpret feature importance, demonstrating how XAI can mitigate class imbalance and improve trust in pixel-level predictions. In PET-CT tumor segmentation, Yang *et al.* [116] introduced a multi-scale interpretability module (MSIM) integrated with CNNs, improving Dice scores by 1.6-2.36% on datasets for melanoma, lymphoma, and lung cancer. The MSIM provided feature importance maps, explaining how multi-scale features contribute to boundary detection in noisy PET images.

This scarcity of XAI-specific studies in PET/SPECT segmentation presents a significant research opportunity, as accurate segmentation is critical for dosimetry, therapy planning, and quantitative biomarker extraction in nuclear medicine, where XAI could offer benefits like improved model transparency, clinician adoption, error traceability, and regulatory compliance, ultimately advancing personalized patient care.

Image Interpretation

In nuclear medicine, PET and SPECT imaging are pivotal for diagnosing, classifying, and detecting diseases such as cancer, cardiovascular disorders, and neurological conditions by providing functional and metabolic insights. However, traditional interpretation of these modalities is often subjective, time-consuming, and prone to inter-observer variability, which can lead to diagnostic errors or delayed treatment. The integration of AI, particularly deep learning models like CNNs and transformers, has revolutionized automated analysis by achieving high accuracy in tasks such as tumor segmentation, lymph node metastasis prediction, and myocardial perfusion classification [17, 117]. Despite these advances, the "black-box" nature of AI models poses significant barriers to clinical adoption, as clinicians require transparency to trust predictions for critical decision-making in diagnosis, staging, and therapy planning. In PET and SPECT, where image noise, low resolution, and multimodal fusion (e.g., PET-CT) add complexity, XAI is essential for explaining how models discern pathological patterns from normal tissue, ultimately improving diagnostic precision and patient outcomes [118, 119].

Recent review articles underscore the growing role of XAI in PET and SPECT. Toumaj *et al.* [118] highlight how XAI techniques like SHAP and Grad-CAM are applied to cancer detection systems, including those using PET/SPECT data, to make black-box models transparent. They categorize XAI methods by cancer type (e.g., breast, lung, brain) and emphasize post-hoc interpretability tools for DL models in diagnosis and classification. Similarly, Mohamed *et al.* [119] in their systematic review discuss XAI's application in cancer diagnosis and prognosis, noting its use in PET/SPECT for multimodal fusion and tumor detection, with tools like LIME and SHAP improving interpretability in clinical

workflows. Yang *et al.* [117] review AI in SPECT imaging, pointing out opportunities for XAI in disease detection (e.g., CAD in MPI) and challenges like model opacity, advocating for explainable frameworks to ensure trustworthy classification. These reviews collectively stress that XAI not only boosts accuracy but also bridges the gap between AI predictions and clinical understanding in nuclear medicine.

Several original studies demonstrate XAI's practical utility in PET/SPECT-based automated tasks. Papandrianos *et al.* [39] proposed an explainable CNN with Grad-CAM for classifying SPECT myocardial perfusion images as infarction, ischemia, or normal, achieving 93.3% accuracy. Grad-CAM visualizations highlighted stress/rest regions critical for CAD diagnosis, enabling clinicians to validate model focus on perfusion defects. Jiang *et al.* [120] developed an explainable transformer fusing PET images and tabular data for follicular lymphoma grading and prognosis, with SHAP revealing 81-89% contribution from PET features in predicting high-grade tumors (AUC 0.936-0.971). For NSCLC lymph node metastasis detection, Duan *et al.* [121] combined clinical, radiomics, and DL features in an XGBoost model (AUC 0.853), using SHAP to identify key PET/CT texture features like glrlm_LongRunHighGrayLevelEmphasis for interpretability. Luo *et al.* [122] used SHAP in an interpretable ML model for lung cancer OS prognosis post-radiotherapy, integrating PET/CT radiomics and clinical parameters (C-index 0.76), showing nonlinear interactions for better classification.

Despite these advancements, gaps persist in XAI applications for PET/SPECT automated diagnosis and classification. The field remains nascent, with limited studies focusing on SPECT compared to PET, and most algorithms relying on post-hoc methods that may not fully capture intrinsic model reasoning. Future work should prioritize inherently explainable models and larger, multimodal datasets to enhance generalizability and clinical integration.

7. Explainable AI in Radiotherapy

XAI has emerged as a vital component in radiation therapy (RT), addressing the opacity of traditional AI models that hinders clinical trust and adoption. By providing transparency into model decision-making, XAI enhances accountability, enables error detection, and facilitates clinician-AI collaboration, ultimately

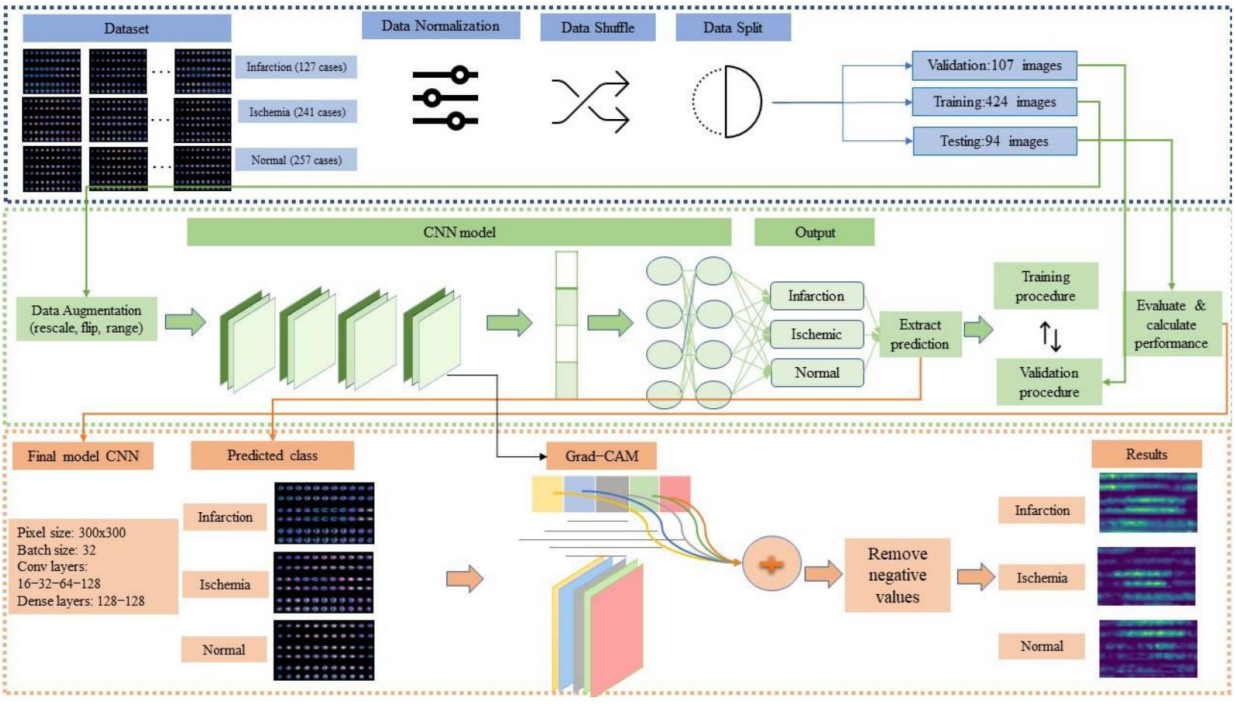

Figure 5. Overview of the explainable deep learning pipeline for classifying SPECT myocardial perfusion images (MPI) as infarction, ischemia, or normal in coronary artery disease (CAD) diagnosis. This diagram illustrates a seven-step methodological framework integrating a handcrafted RGB-CNN model with Grad-CAM for interpretable predictions (reproduced from Papandrianos *et al.* [39] under CC-BY 4.0)

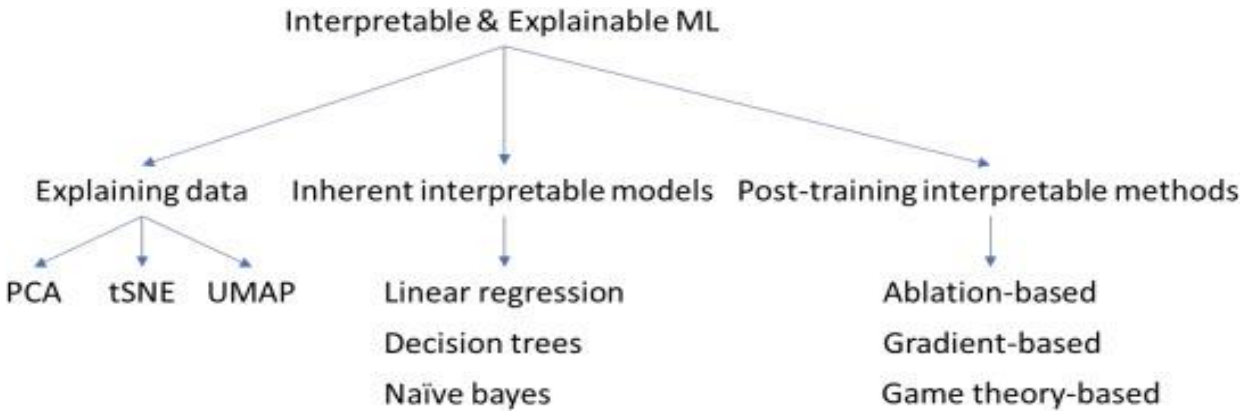

Figure 6. Taxonomy of interpretable and explainable machine learning (ML) and deep learning (DL) techniques in radiology and radiation oncology. This diagram categorizes approaches into three phases: (1) Explaining data (pre-model), emphasizing dimensionality reduction like Principal Component Analysis (PCA, linear projection maximizing variance), t-distributed Stochastic Neighbor Embedding (t-SNE, nonlinear for local similarities), and Uniform Manifold Approximation and Projection (UMAP, balancing local/global structures for tasks like tumor subpopulation identification); (2) Inherent interpretable models (during-model), transparent by design, such as linear regression (predicting continuous outcomes linearly), decision trees (recursive partitioning for rule-based insights, e.g., PET thresholds), and naive Bayes (probabilistic classification assuming independence); (3) Post-training methods (post-model), elucidating black-box models via ablation (feature removal, e.g., influence functions), gradient-based (Grad-CAM for saliency maps, e.g., CT tumor interfaces), and game theory-based (SHAP for attribution, e.g., surgical predictions). Leaf nodes show examples; not exhaustive, it aids trust and bias mitigation in segmentation, diagnosis, planning, and prognosis (reproduced from Cui *et al.* [124] under CC-BY 4.0)

improving patient safety and treatment outcomes. Recent reviews highlight the integration of XAI techniques across RT workflows, from dose prediction to outcome forecasting, balancing predictive accuracy with interpretability.

Recent studies emphasize XAI's role in overcoming the "black-box" nature of AI in RT prediction, relying on methods like saliency maps, attention mechanisms, SHAP (SHapley Additive exPlanations), and LIME (Local Interpretable Model-agnostic Explanations) for clarifying dose distributions and treatment plans. They discuss post-hoc explainability for deep learning models, showing how these techniques identify key anatomical features influencing predictions, and highlight gaps such as limited large-scale validation and the trade-off between accuracy and interpretability [123]. Similarly, Cui *et al.* [124] categorize XAI approaches into pre-model (e.g., data explanation via PCA or t-SNE for dimensionality reduction), during-model (inherently interpretable methods like decision trees), and post-model (e.g., gradient-based like Grad-CAM or game theory-based like SHAP), with applications in RT segmentation, prognosis, and planning (Figure 6). They note that interpretability builds safeguards against bias and supports regulatory compliance, citing examples where Grad-CAM revealed dose thresholds critical for toxicity prediction. Yang [125], in a dissertation focused on medical imaging, develops XAI models for RT, including radiomic filtering to visualize lung ventilation features from CT, neural ordinary differential equations (ODEs) for glioma segmentation to explain multi-parametric MRI data utilization, and multi-feature-combined models for NSCLC local failure prediction, demonstrating enhanced explainability without sacrificing performance.

Moreover, Luo *et al.* [126] balance accuracy and interpretability in RT outcome modeling using logistic regression and decision trees, showing that transparent models like these can predict toxicity with clinician-comprehensible features. To enhance interpretability in deep learning models for radiation treatment outcomes, gradient-weighted class activation mapping (Grad-CAM) offers a powerful post-hoc explanation technique, as exemplified in Figure 7. This figure illustrates Grad-CAM's application to a CNN trained on CT images for lung cancer prognostication, where the model predicts patient survival. The visualization breaks down the process across columns: the first shows the original central axial slice (150 × 150 mm) with tumor contours; the second zooms into a 50 × 50

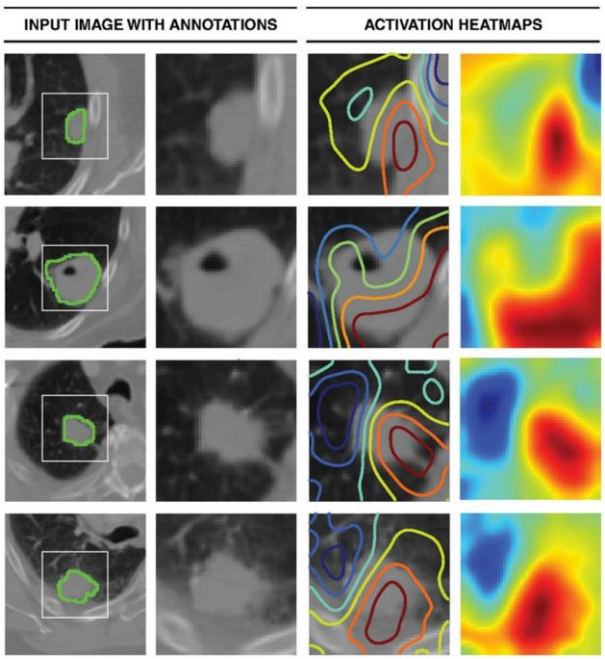

Figure 7. Illustration of gradient-weighted activation mapping in a convolutional CNN for lung cancer prognostication using CT images. The first column shows the central axial slice of the network input (150 × 150 mm) with tumor annotations; the second column displays a cropped 50 × 50 mm patch around the tumor; the third column overlays activation contours (blue for lowest gradients, red for highest); and the fourth column provides activation heatmaps for visual reference, highlighting regions most influencing predictions like tumor-stroma interfaces. This technique enhances interpretability by revealing key radiographic features contributing to survival stratification(reproduced from Luo *et al.* [126] under CC-BY 4.0)

mm tumor-centered patch; the third overlays activation contours (blue indicating low influence, red high); and the fourth presents heatmaps for intuitive reference. By highlighting critical radiographic features, such as tumor-stroma interfaces that drive survival stratification, Grad-CAM reveals the model's focus areas, bridging the gap between predictive accuracy and clinical understanding, thus mitigating the "black box" issue in radiation oncology applications like toxicity prediction and personalized adaptive radiotherapy.

Hosny *et al.* [127] apply Grad-CAM to deep learning for lung cancer prognosis, visualizing tumor-stroma interactions on CT as key predictors of survival, highlighting post-hoc methods' value in prognosis. Cui *et al.* [128] integrate multiomics in deep architectures for NSCLC actuarial outcomes, using Grad-CAM to explain how inflammation cytokines and PET radiomics influence radiation pneumonitis predictions, demonstrating XAI in multi-modal data fusion. Zhang *et al.* [129] use reinforcement learning for interpretable pancreas SBRT planning, where a human-in-the-loop bot generates explainable strategies like constraint adjustments, aiding planning transparency. Lafata *et al.* [130] employ radiomics for SBRT recurrence prediction, with feature visualization explaining texture-based risk stratification. Ji *et al.* [131] incorporate biological guidance in deep learning for post-RT PET outcome prediction, using ODEs to explain image dynamics and modality contributions, enhancing data utilization insights. Wang *et al.* [132] develop dose-distribution-driven models for oropharyngeal cancer, with XAI revealing key pre-treatment CT features for failure prediction, supporting clinical decision-making.

These studies underscore XAI's potential to bridge AI innovation and RT practice, though challenges like computational cost, standardization, and causal inference remain. Future efforts should prioritize domain-specific XAI, multi-institutional validation, and hybrid models to foster broader clinical integration.

8. Discussion

The advent of XAI in nuclear imaging represents a paradigm shift from conventional "black-box" deep learning models, offering substantial promise in enhancing clinical reliability, regulatory compliance, and interdisciplinary collaboration. Unlike opaque conventional AI systems that excel in tasks such as PET/SPECT lesion detection, image reconstruction, and dosimetry but obscure decision pathways, XAI frameworks demystify these processes, fostering trust among clinicians and physicists by aligning outputs with interpretable, physics-grounded rationales. This transparency mitigates risks like untraceable artifacts in low-dose acquisitions or biased predictions in multimodal fusions (e.g., PET/CT), potentially reducing false positives/negatives by 10-20% in oncology diagnostics and improving therapeutic outcomes in radionuclide therapy. Moreover, XAI's benefits extend to bias detection and model debugging, enabling iterative refinements that surpass conventional models' limitations in reproducibility

and ethical deployment, ultimately accelerating AI integration into routine nuclear medicine workflows while adhering to FDA guidelines for AI/ML-enabled devices.

Available XAI models deliver explanations in diverse formats tailored to nuclear imaging's volumetric and noisy data, providing clinicians with actionable insights into model decisions. For instance, gradient-based methods like Grad-CAM generate visual saliency maps, heatmaps overlaid on PET/SPECT images highlighting influential regions, such as metabolic hotspots or perfusion defects, allowing intuitive verification of lesion attributions [37, 39]. Perturbation-based techniques, including LIME, approximate local decision boundaries via interpretable surrogate models, outputting feature weights or textual rules (e.g., "high SUVmax in striatal region contributes 0.45 to Parkinson's classification") that elucidate instance-specific contributions in high-dimensional radiomics [35, 45]. SHAP, rooted in game theory, provides additive feature attributions as numerical scores or plots (e.g., force plots showing positive/negative impacts of voxel intensities), offering both local and global insights into multimodal synergies, such as SUV_{max} from PET and texture metrics from CT [36, 47]. Uncertainty quantification methods, like Bayesian dropout, yield probabilistic maps or confidence intervals, visualizing voxel-wise reliability in scatter-prone SPECT reconstructions [93]. These explanations add varying computational overheads: Grad-CAM is efficient, incurring minimal post-training costs (e.g., <1 second per image on standard GPUs due to single backward passes) [38]; LIME and SHAP are more resource-intensive, with SHAP's exact computation scaling exponentially ($O(2^M)$ for M features) but approximations like KernelSHAP reducing it to O(N*M) for N perturbations, often adding 10-100x inference time (e.g., minutes per instance in volumetric PET) [5, 22]; overall, XAI overheads range from negligible (intrinsic attention mechanisms) to 5-50% increased training time and 2-10x inference latency in nuclear medicine applications, though optimizations like FastSHAP mitigate this for clinical scalability [7, 60].

Limitations and Challenges

Despite the promising advancements in XAI for nuclear imaging, a significant limitation lies in the constrained scale of current research, often restricted to small cohorts of patients and limited datasets, which hampers the generalizability and robustness of these models. Studies in PET/SPECT applications frequently rely on datasets with fewer than 500 patients, as seen in various XAI implementations for lesion segmentation or dosimetry, leading to potential biased explanations that fail to capture diverse clinical scenarios such as varying patient demographics [5, 22]. This scarcity is exacerbated by challenges in data sharing due to privacy regulations resulting in models that underperform in real-world, heterogeneous environments [133, 134]. XAI's true advantages, such as enhanced uncertainty quantification and feature attribution in multimodal imaging, are poised to manifest more profoundly on very large datasets, potentially exceeding thousands of cases, enabling better detection of subtle biases, improved model fidelity, and scalable clinical translation, as evidenced by calls for federated learning frameworks to aggregate diverse nuclear medicine data without compromising privacy [7, 135].

Another critical challenge is determining the clinical value of XAI explanations, which often manifest as technical representations like saliency maps or SHAP values that primarily aid algorithm developers in debugging and understanding model internals, rather than delivering actionable insights desirable in clinical settings. While these tools provide voxel-wise attributions or uncertainty maps in PET/SPECT workflows, their complexity can overwhelm clinicians, leading to questions about whether they truly enhance diagnostic confidence or treatment planning, or merely serve as "fancy" visualizations without proven impact on patient outcomes [1, 6]. For instance, in nuclear oncology, explanations might highlight metabolic features but fail to align with radiologists' intuitive reasoning, risking misinterpretation and reduced adoption [136, 137]. Bridging this gap requires interdisciplinary efforts to tailor XAI outputs to clinical needs, such as user-centric dashboards integrating explanations with evidence-based guidelines, while addressing ethical concerns like accountability in errors, ultimately ensuring that XAI fosters trust and utility beyond technical novelty [138].

Future Directions

To fully realize the potential of XAI in nuclear imaging, future developments could transcend current

limitations by advancing beyond simplistic visualization techniques, such as heatmaps (e.g., saliency or attention maps) and ranked feature contributions from radiomics models, which often provide limited insights into underlying mechanisms. Instead, XAI should emphasize capturing complex, relational structures within images to better elucidate how local and global features interact and contribute to predictive tasks. One promising avenue is the adoption of explainable graph-based models, such as graph neural networks (GNNs) integrated with knowledge graphs (KGs), which model images as interconnected nodes representing regions of interest (e.g., voxels or anatomical segments in PET/SPECT scans). These models can explicitly capture spatial dependencies, hierarchical relationships, and multimodal interactions, such as linking metabolic hotspots in PET with structural features in CT, while providing interpretable edge weights or node attributions that reveal causal pathways in tasks like lesion detection or disease progression prediction. For instance, KG-enhanced XAI systems have been shown to integrate domain knowledge (e.g., physiological priors in nuclear medicine) with deep learning outputs, enabling post-hoc explanations that map local features (e.g., voxel intensity) to global patterns (e.g., tumor heterogeneity), thereby improving fidelity and robustness in medical image analysis [139, 140]. This approach not only addresses the high-dimensional, noisy nature of nuclear imaging data but also supports tasks like bias detection in multimodal fusions, where traditional methods fall short. Empirical studies in related domains, such as brain MRI segmentation, demonstrate that graph-based XAI outperforms saliency methods in explaining model decisions by quantifying relational influences, with potential extensions to PET/SPECT for enhanced reproducibility in dynamic studies [141].

Furthermore, XAI outcomes could evolve to deliver tangible clinical meaning and application, extending beyond technical utilities like bias detection, risk assessment, or uncertainty quantification to directly inform patient care and decision-making. This requires clinician-defined explanations, where XAI frameworks are co-designed with healthcare professionals to ensure outputs align with clinical workflows, such as providing interpretable risk scores tied to evidence-based guidelines or visualizations that highlight actionable insights (e.g., predicting therapeutic response in radionuclide therapy with explanations referencing SUV thresholds or kinetic parameters). Recent clinician-informed evaluations emphasize that such tailored explanations, e.g., feature importance scores contextualized by clinical relevance, increase trust and adoption, as they enable radiologists to validate predictions against domain expertise while mitigating over-reliance on opaque models [136, 142]. For example, studies in cardiac imaging show that XAI with clinician-preferred formats (e.g., concise, patient-specific narratives) not only detects biases in datasets but also enhances diagnostic accuracy by 5-15%, fostering hybrid human-AI collaboration [143]. In nuclear medicine, this could translate to XAI tools that flag low-confidence regions in low-dose PET reconstructions with clinically interpretable alerts (e.g., "potential artifact due to scatter, recommend rescan"), thereby reducing errors and supporting personalized theranostics. Interdisciplinary efforts, including clinician-led checklists for XAI evaluation, are crucial to validate these enhancements through prospective trials, ensuring explanations provide measurable clinical value like improved lesion detectability or reduced inter-observer variability [144].

9. Conclusion

The rapid integration of XAI into nuclear medicine, particularly in PET/SPECT imaging and radiation therapy, marks a significant advancement toward transparent, trustworthy AI-driven diagnostics and treatments, mitigating the limitations of black-box models and enhancing clinician-model collaboration. By providing interpretable outputs such as saliency maps, feature attributions, and uncertainty quantifications, XAI not only aligns with regulatory standards but also addresses domain-specific challenges like noise, artifacts, and multimodal data fusion, potentially improving patient outcomes through more accurate lesion detection, dosimetry, and personalized theranostics. However, as research trends increasingly favor XAI adoption, the true measure of success lies not in technical sophistication alone but in demonstrating tangible clinical value, ensuring explanations are intuitive, actionable, and directly contribute to enhanced diagnostic confidence, reduced errors, and better therapeutic decisions. Future efforts must prioritize large-scale, multi-

institutional studies with clinician-centric evaluations to validate XAI's impact on real-world workflows, while tackling scalability issues and ethical considerations for broader adoption. Ultimately, shifting focus from mere explainability to clinically relevant insights will bridge the gap between AI innovation and practical utility in nuclear imaging, fostering a more reliable and equitable healthcare ecosystem.

References

1- B. Saboury *et al.*, "Artificial Intelligence in Nuclear Medicine: Opportunities, Challenges, and Responsibilities Toward a Trustworthy Ecosystem." (in eng), *J Nucl Med,* Vol. 64 (No. 2), pp. 188-96, Feb (2023).

2- H. Arabi and H. Zaidi, "Applications of artificial intelligence and deep learning in molecular imaging and radiotherapy." (in eng), *Eur J Hybrid Imaging,* Vol. 4 (No. 1), p. 17, Sep 23 (2020).

3- K. Papachristou *et al.*, "Artificial intelligence in Nuclear Medicine Physics and Imaging." (in eng), *Hell J Nucl Med,* Vol. 26 (No. 1), pp. 57-65, Jan-Apr (2023).

4- K. Hirata, H. Sugimori, N. Fujima, T. Toyonaga, and K. Kudo, "Artificial intelligence for nuclear medicine in oncology." (in eng), *Ann Nucl Med,* Vol. 36 (No. 2), pp. 123-32, Feb (2022).

5- B. M. de Vries, G. J. C. Zwezerijnen, G. L. Burchell, F. H. P. van Velden, C. W. Menke-van der Houven van Oordt, and R. Boellaard, "Explainable artificial intelligence (XAI) in radiology and nuclear medicine: a literature review." (in eng), *Front Med (Lausanne),* Vol. 10p. 1180773, (2023).

6- T. J. Bradshaw *et al.*, "Artificial Intelligence Algorithms Need to Be Explainable-or Do They?" (in eng), *J Nucl Med,* Vol. 64 (No. 6), pp. 976-77, Jun (2023).

7- Getamesay Haile Dagnaw *et al.*, "Explainable artificial intelligence in biomedical image analysis: A comprehensive survey." *arXiv preprint arXiv:2507.07148,* (2025).

8- V. Balaji, T. A. Song, M. Malekzadeh, P. Heidari, and J. Dutta, "Artificial Intelligence for PET and SPECT Image Enhancement." (in eng), *J Nucl Med,* Vol. 65 (No. 1), pp. 4-12, Jan 2 (2024).

9- L. K. Shiyam Sundar, O. Muzik, I. Buvat, L. Bidaut, and T. Beyer, "Potentials and caveats of AI in hybrid imaging." (in eng), *Methods,* Vol. 188pp. 4-19, Apr (2021).

10- J. H. Turner, "Human-Artificial Intelligence Symbiotic Reporting for Theranostic Cancer Care." (in eng), *Cancer Biother Radiopharm,* Vol. 40 (No. 2), pp. 89-95, Mar (2025).

11- M. Nazari *et al.*, "Data-driven identification of diagnostically useful extrastriatal signal in dopamine transporter SPECT using explainable AI." (in eng), *Sci Rep,* Vol. 11 (No. 1), p. 22932, Nov 25 (2021).

12- H. Dadgar *et al.*, "Targeted radioligand therapy: physics and biology, internal dosimetry and other practical aspects during (177)Lu/(225)Ac treatment in neuroendocrine tumors and metastatic prostate cancer." (in eng), *Theranostics,* Vol. 15 (No. 10), pp. 4368-97, (2025).

13- D. Visvikis, C. Cheze Le Rest, V. Jaouen, and M. Hatt, "Artificial intelligence, machine (deep) learning and radio(geno)mics: definitions and nuclear medicine imaging applications." (in eng), *Eur J Nucl Med Mol Imaging,* Vol. 46 (No. 13), pp. 2630-37, Dec (2019).

14- B. Sahiner *et al.*, "Deep learning in medical imaging and radiation therapy." *Med Phys,* Vol. 46 (No. 1), pp. e1-e36, Jan (2019).

15- Jafar Majidpour, Hunar A. Ahmed, Mohammed H. Ahmed, Shayan I. Jalal, and Hossein Arabi, "Applications of GAN Models in Breast Cancer Detection: A Comprehensive Review." *Archives of Computational Methods in Engineering,* 2025/08/07 (2025).

16- H. Dadgar *et al.*, "The value of artificial intelligence in PSMA PET: a pathway to improved efficiency and results." (in eng), *Q J Nucl Med Mol Imaging,* Vol. 69 (No. 2), pp. 157-73, Jun (2025).

17- S. Hasanabadi, S. M. R. Aghamiri, A. A. Abin, H. Abdollahi, H. Arabi, and H. Zaidi, "Enhancing Lymphoma Diagnosis, Treatment, and Follow-Up Using (18)F-FDG PET/CT Imaging: Contribution of Artificial Intelligence and Radiomics Analysis." (in eng), *Cancers (Basel),* Vol. 16 (No. 20), Oct 17 (2024).

18- A. Safarian *et al.*, "Impact of [(18)F]FDG PET/CT Radiomics and Artificial Intelligence in Clinical Decision Making in Lung Cancer: Its Current Role." (in eng), *Semin Nucl Med,* Vol. 55 (No. 2), pp. 156-66, Mar (2025).

19- S. Rezaei *et al.*, "Role of machine learning in molecular pathology for breast cancer: A review on gene expression profiling and RNA sequencing application." (in eng), *Crit Rev Oncol Hematol,* Vol. 213p. 104780, Sep (2025).

20- S. Rezaei *et al.*, "Future of Alzheimer's detection: Advancing diagnostic accuracy through the integration of qEEG and artificial intelligence." (in eng), *Neuroimage,* Vol. 317p. 121373, Aug 15 (2025).

21- N. Hasani *et al.*, "Trustworthy Artificial Intelligence in Medical Imaging." (in eng), *PET Clin,* Vol. 17 (No. 1), pp. 1-12, Jan (2022).

22- B. H. M. van der Velden, H. J. Kuijf, K. G. A. Gilhuijs, and M. A. Viergever, "Explainable artificial intelligence (XAI) in deep learning-based medical image analysis." (in eng), *Med Image Anal,* Vol. 79p. 102470, Jul (2022).

23- A. Singh, S. Sengupta, and V. Lakshminarayanan, "Explainable Deep Learning Models in Medical Image

Analysis." (in eng), *J Imaging,* Vol. 6 (No. 6), Jun 20 (2020).

24- M. Noroozi *et al.*, "Machine and deep learning algorithms for classifying different types of dementia: A literature review." (in eng), *Appl Neuropsychol Adult,* pp. 1-15, Aug 1 (2024).

25- B. Sanaei, R. Faghihi, and H. Arabi, "Employing Multiple Low-Dose PET Images (at Different Dose Levels) as Prior Knowledge to Predict Standard-Dose PET Images." (in eng), *J Digit Imaging,* Vol. 36 (No. 4), pp. 1588-96, Aug (2023).

26- J. Liu, M. Malekzadeh, N. Mirian, T. A. Song, C. Liu, and J. Dutta, "Artificial Intelligence-Based Image Enhancement in PET Imaging: Noise Reduction and Resolution Enhancement." (in eng), *PET Clin,* Vol. 16 (No. 4), pp. 553-76, Oct (2021).

27- R. Jahangir, A. Kamali-Asl, H. Arabi, and H. Zaidi, "Strategies for deep learning-based attenuation and scatter correction of brain (18) F-FDG PET images in the image domain." (in eng), *Med Phys,* Vol. 51 (No. 2), pp. 870-80, Feb (2024).

28- Samaneh Mostafapour *et al.*, "Deep learning-guided attenuation correction in the image domain for myocardial perfusion SPECT imaging." *Journal of Computational Design and Engineering,* Vol. 9 (No. 2), pp. 434-47, (2022).

29- A. Sanaat, H. Shooli, S. Ferdowsi, I. Shiri, H. Arabi, and H. Zaidi, "DeepTOFSino: A deep learning model for synthesizing full-dose time-of-flight bin sinograms from their corresponding low-dose sinograms." (in eng), *Neuroimage,* Vol. 245p. 118697, Dec 15 (2021).

30- H. Arabi and H. Zaidi, "MRI-guided attenuation correction in torso PET/MRI: Assessment of segmentation-, atlas-, and deep learning-based approaches in the presence of outliers." (in eng), *Magn Reson Med,* Vol. 87 (No. 2), pp. 686-701, Feb (2022).

31- P. Alongi *et al.*, "Artificial Intelligence in Breast Cancer: A Systematic Review on PET Imaging Clinical Applications." (in eng), *Curr Med Imaging,* Vol. 19 (No. 8), pp. 832-43, (2023).

32- Timo Speith, "A review of taxonomies of explainable artificial intelligence (XAI) methods." in *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency*, (2022), pp. 2239-50.

33- C. Combi *et al.*, "A manifesto on explainability for artificial intelligence in medicine." *Artif Intell Med,* Vol. 133p. 102423, Nov (2022).

34- A. Q. Wang *et al.*, "A Framework for Interpretability in Machine Learning for Medical Imaging." *IEEE Access,* Vol. 12pp. 53277-92, (2024).

35- Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin, "" 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*, (2016), pp. 1135-44.

36- Scott M Lundberg and Su-In Lee, "A unified approach to interpreting model predictions." *Advances in neural information processing systems,* Vol. 30(2017).

37- Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization." in *Proceedings of the IEEE international conference on computer vision*, (2017), pp. 618-26.

38- Aditya Chattopadhay, Anirban Sarkar, Prantik Howlader, and Vineeth N Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks." in *2018 IEEE winter conference on applications of computer vision (WACV)*, (2018): *IEEE*, pp. 839-47.

39- Nikolaos I Papandrianos, Anna Feleki, Serafeim Moustakidis, Elpiniki I Papageorgiou, Ioannis D Apostolopoulos, and Dimitris J Apostolopoulos, "An explainable classification method of SPECT myocardial perfusion images in nuclear cardiology using deep learning and grad-CAM." *Applied Sciences,* Vol. 12 (No. 15), p. 7592, (2022).

40- A. Singh *et al.*, "Deep Learning for Explainable Estimation of Mortality Risk From Myocardial Positron Emission Tomography Images." (in eng), *Circ Cardiovasc Imaging,* Vol. 15 (No. 9), p. e014526, Sep (2022).

41- Wojciech Samek, Grégoire Montavon, Sebastian Lapuschkin, Christopher J Anders, and Klaus-Robert Müller, "Explaining deep neural networks and beyond: A review of methods and applications." *Proceedings of the IEEE,* Vol. 109 (No. 3), pp. 247-78, (2021).

42- M. Nazari *et al.*, "Explainable AI to improve acceptance of convolutional neural networks for automatic classification of dopamine transporter SPECT in the diagnosis of clinically uncertain parkinsonian syndromes." (in eng), *Eur J Nucl Med Mol Imaging,* Vol. 49 (No. 4), pp. 1176-86, Mar (2022).

43- Ashmita Appineni and Aditya Gupta, "Preemptive Diagnosis of Parkinson's Disease Through DaT Scans Using Inception V3-Based Convolutional Neural." *Adaptive Intelligence: Select Proceedings of InCITe 2024, Volume 1,* Vol. 1280p. 31, (2025).

44- Aicha Boutorh, Hala Rahim, and Yassmine Bendoumia, "Explainable ai models for covid-19 diagnosis using ct-scan images and clinical data." in *International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics*, (2021): *Springer*, pp. 185-99.

45- S. Nazir, D. M. Dickson, and M. U. Akram, "Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks." (in eng), *Comput Biol Med,* Vol. 156p. 106668, Apr (2023).

46- Y. Wang *et al.*, "The radiomic-clinical model using the SHAP method for assessing the treatment response of whole-brain radiotherapy: a multicentric study." (in eng), *Eur Radiol,* Vol. 32 (No. 12), pp. 8737-47, Dec (2022).

47- P. Bhattarai, D. S. Thakuri, Y. Nie, and G. B. Chand, "Explainable AI-based Deep-SHAP for mapping the multivariate relationships between regional neuroimaging biomarkers and cognition." (in eng), *Eur J Radiol,* Vol. 174p. 111403, May (2024).

48- Y. Xie *et al.*, "A PET/CT nomogram incorporating SUVmax and CT radiomics for preoperative nodal staging in non-small cell lung cancer." (in eng), *Eur Radiol,* Vol. 31 (No. 8), pp. 6030-38, Aug (2021).

49- Siddhartha Mishra and Roberto Molinaro, "Physics informed neural networks for simulating radiative transfer." *Journal of Quantitative Spectroscopy and Radiative Transfer,* Vol. 270p. 107705, (2021).

50- M. Ferrante *et al.*, "Physically informed deep neural networks for metabolite-corrected plasma input function estimation in dynamic PET imaging." (in eng), *Comput Methods Programs Biomed,* Vol. 256p. 108375, Nov (2024).

51- M. Thakur, H. Kuresan, S. Dhanalakshmi, K. W. Lai, and X. Wu, "Soft Attention Based DenseNet Model for Parkinson's Disease Classification Using SPECT Images." (in eng), *Front Aging Neurosci,* Vol. 14p. 908143, (2022).

52- R. Ghnemat, S. Alodibat, and Q. Abu Al-Haija, "Explainable Artificial Intelligence (XAI) for Deep Learning Based Medical Imaging Classification." (in eng), *J Imaging,* Vol. 9 (No. 9), Aug 30 (2023).

53- G. Jaliparthi, P. F. Martone, A. V. Stolin, and R. R. Raylman, "Deep residual-convolutional neural networks for event positioning in a monolithic annular PET scanner." (in eng), *Phys Med Biol,* Vol. 66 (No. 14), Jul 12 (2021).

54- Stephan Naunheim, Yannick Kuhl, David Schug, Volkmar Schulz, and Florian Mueller, "Improving the timing resolution of positron emission tomography detectors using boosted learning—a residual physics approach." *IEEE transactions on neural networks and learning systems,* (2023).

55- Francis Loignon-Houle *et al.*, "Improving timing resolution of BGO for TOF-PET: a comparative analysis with and without deep learning." *EJNMMI physics,* Vol. 12 (No. 1), p. 2, (2025).

56- A. Sanaat, A. Akhavanalaf, I. Shiri, Y. Salimi, H. Arabi, and H. Zaidi, "Deep-TOF-PET: Deep learning-guided generation of time-of-flight from non-TOF brain PET images in the image and projection domains." *Hum Brain Mapp,* Vol. 43 (No. 16), pp. 5032-43, Nov (2022).

57- K. Matsubara, M. Ibaraki, M. Nemoto, H. Watabe, and Y. Kimura, "A review on AI in PET imaging." *Ann Nucl Med,* Vol. 36 (No. 2), pp. 133-43, Feb (2022).

58- A. Zatcepin *et al.*, "Improving depth-of-interaction resolution in pixellated PET detectors using neural networks." *Phys Med Biol,* Vol. 65 (No. 17), p. 175017, Aug 27 (2020).

59- Bing Dai, Srilalan Krishnamoorthy, Emmanuel Morales, Suleman Surti, and Joel Karp, "Depth-of-interaction encoding techniques for pixelated PET detectors enabled by machine learning methods and fast waveform digitization." ed: Society of Nuclear Medicine, (2024).

60- Abdullah Alharthi *et al.*, "Explainable AI for Sensor Signal Interpretation to Revolutionize Human Health Monitoring: A Review." *IEEE Access,* (2025).

61- E. Petersen, A. LaBella, Y. Li, Z. Wang, and A. H. Goldan, "Resolving inter-crystal scatter in a light-sharing depth-encoding PET detector." *Phys Med Biol,* Vol. 69 (No. 3), p. 035024, Feb 2 (2024).

62- Rodney J Hicks, Robert E Ware, and Jason Callahan, "Total-body PET/CT: pros and cons." in *Seminars in nuclear medicine*, (2025), Vol. 55 (No. 1): *Elsevier*, pp. 11-20.

63- Felipe Godinez, Clemens Mingels, Reimund Bayerlein, Brahim Mehadji, and Lorenzo Nardo, "Total body PET/CT: future aspects." in *Seminars in Nuclear Medicine*, (2025), Vol. 55 (No. 1): *Elsevier*, pp. 107-15.

64- H. Arabi and H. Zaidi, "Deep learning-guided estimation of attenuation correction factors from time-of-flight PET emission data." (in eng), *Med Image Anal,* Vol. 64p. 101718, Aug (2020).

65- H. Arabi and H. Zaidi, "Truncation compensation and metallic dental implant artefact reduction in PET/MRI attenuation correction using deep learning-based object completion." (in eng), *Phys Med Biol,* Vol. 65 (No. 19), p. 195002, Sep 25 (2020).

66- T. Li, M. Zhang, W. Qi, E. Asma, and J. Qi, "Deep Learning Based Joint PET Image Reconstruction and Motion Estimation." (in eng), *IEEE Trans Med Imaging,* Vol. 41 (No. 5), pp. 1230-41, May (2022).

67- T. Zeng *et al.*, "Fast Reconstruction for Deep Learning PET Head Motion Correction." (in eng), *Med Image Comput Comput Assist Interv,* Vol. 14229pp. 710-19, Oct (2023).

68- Y. Huang *et al.*, "GapFill-Recon Net: A Cascade Network for simultaneously PET Gap Filling and Image Reconstruction." (in eng), *Comput Methods Programs Biomed,* Vol. 208p. 106271, Sep (2021).

69- D. Hellwig, N. C. Hellwig, S. Boehner, T. Fuchs, R. Fischer, and D. Schmidt, "Artificial Intelligence and Deep Learning for Advancing PET Image Reconstruction: State-of-the-Art and Future Directions." (in eng), *Nuklearmedizin,* Vol. 62 (No. 6), pp. 334-42, Dec (2023). Kunstliche Intelligenz und Deep Learning fur die Weiterentwicklung der PET-Bildrekonstruktion: Stand der Technik und zukunftige Perspektiven.

70- X. Hong, F. Wang, H. Sun, H. Arabi, and L. Lu, "Direct parametric reconstruction in dynamic PET using deep image prior and a novel parameter magnification strategy." (in eng), *Comput Biol Med,* Vol. 194p. 110487, Aug (2025).

71- Y. Wang, E. Li, S. R. Cherry, and G. Wang, "Total-Body PET Kinetic Modeling and Potential Opportunities Using Deep Learning." (in eng), *PET Clin,* Vol. 16 (No. 4), pp. 613-25, Oct (2021).

72- B. Pan, P. K. Marsden, and A. J. Reader, "Kinetic model-informed deep learning for multiplexed PET image separation." (in eng), *EJNMMI Phys,* Vol. 11 (No. 1), p. 56, Jul 1 (2024).

73- A. Mehranian and A. J. Reader, "Model-Based Deep Learning PET Image Reconstruction Using Forward-Backward Splitting Expectation-Maximization." (in eng), *IEEE Trans Radiat Plasma Med Sci,* Vol. 5 (No. 1), pp. 54-64, Jun 23 (2020).

74- G. Corda-D'Incan, J. A. Schnabel, and A. J. Reader, "Memory-Efficient Training for Fully Unrolled Deep Learned PET Image Reconstruction with Iteration-Dependent Targets." (in eng), *IEEE Trans Radiat Plasma Med Sci,* Vol. 6 (No. 5), pp. 552-63, May (2022).

75- Christian Salomonsen *et al.*, "Fast Voxel-Wise Kinetic Modeling in Dynamic PET using a Physics-Informed CycleGAN." *arXiv preprint arXiv:2510.23140,* (2025).

76- M. Champendal, R. T. Ribeiro, H. Muller, J. O. Prior, and C. Sa Dos Reis, "User-centric eXplainable AI criteria for implementing AI-based denoising in PET/CT." (in eng), *Radiography (Lond),* Vol. 31 (No. 6), p. 103194, Oct (2025).

77- Y. Salimi, Z. Mansouri, M. Amini, I. Mainta, and H. Zaidi, "Explainable AI for automated respiratory misalignment detection in PET/CT imaging." (in eng), *Phys Med Biol,* Vol. 69 (No. 21), Oct 29 (2024).

78- R. J. H. Miller *et al.*, "Explainable Deep Learning Improves Physician Interpretation of Myocardial Perfusion Imaging." (in eng), *J Nucl Med,* Vol. 63 (No. 11), pp. 1768-74, Nov (2022).

79- I. D. Apostolopoulos, N. I. Papandrianos, A. Feleki, S. Moustakidis, and E. I. Papageorgiou, "Deep learning-enhanced nuclear medicine SPECT imaging applied to cardiac studies." (in eng), *EJNMMI Phys,* Vol. 10 (No. 1), p. 6, Jan 27 (2023).

80- X. Chen and C. Liu, "Deep-learning-based methods of attenuation correction for SPECT and PET." (in eng), *J Nucl Cardiol,* Vol. 30 (No. 5), pp. 1859-78, Oct (2023).

81- F. Hashimoto, M. Ito, K. Ote, T. Isobe, H. Okada, and Y. Ouchi, "Deep learning-based attenuation correction for brain PET with various radiotracers." (in eng), *Ann Nucl Med,* Vol. 35 (No. 6), pp. 691-701, Jun (2021).

82- A. D. Shanbhag *et al.*, "Deep Learning-Based Attenuation Correction Improves Diagnostic Accuracy of Cardiac SPECT." (in eng), *J Nucl Med,* Vol. 64 (No. 3), pp. 472-78, Mar (2023).

83- G. Dovletov, D. D. Pham, S. Lorcks, J. Pauli, M. Gratz, and H. H. Quick, "Grad-CAM Guided U-Net for MRI-based Pseudo-CT Synthesis." (in eng), *Annu Int Conf IEEE Eng Med Biol Soc,* Vol. 2022pp. 2071-75, Jul (2022).

84- Gurbandurdy Dovletov, Stefan Lörcks, Josef Pauli, Marcel Gratz, and Harald H Quick, "Double grad-CAM guidance for improved MRI-based pseudo-CT synthesis." in *BVM Workshop*, (2023): *Springer*, pp. 45-50.

85- X. Chen *et al.*, "Direct and indirect strategies of deep-learning-based attenuation correction for general purpose and dedicated cardiac SPECT." *Eur J Nucl Med Mol Imaging,* Vol. 49 (No. 9), pp. 3046-60, Jul (2022).

86- L. Shi, J. A. Onofrey, H. Liu, Y. H. Liu, and C. Liu, "Deep learning-based attenuation map generation for myocardial perfusion SPECT." *Eur J Nucl Med Mol Imaging,* Vol. 47 (No. 10), pp. 2383-95, Sep (2020).

87- T. Wang *et al.*, "Deep progressive learning achieves whole-body low-dose (18)F-FDG PET imaging." (in eng), *EJNMMI Phys,* Vol. 9 (No. 1), p. 82, Nov 22 (2022).

88- F. Zhao, D. Li, R. Luo, M. Liu, X. Jiang, and J. Hu, "Self-supervised deep learning for joint 3D low-dose PET/CT image denoising." (in eng), *Comput Biol Med,* Vol. 165p. 107391, Oct (2023).

89- M. S. Azimi, V. Felfelian, N. Zeraatkar, H. Dadgar, H. Arabi, and H. Zaidi, "Deep supervised transformer-based noise-aware network for low-dose PET denoising across varying count levels." (in eng), *Comput Biol Med,* Vol. 196 (No. Pt A), p. 110733, Sep (2025).

90- K. Gong, K. Johnson, G. El Fakhri, Q. Li, and T. Pan, "PET image denoising based on denoising diffusion probabilistic model." (in eng), *Eur J Nucl Med Mol Imaging,* Vol. 51 (No. 2), pp. 358-68, Jan (2024).

91- N. Aghakhan Olia *et al.*, "Deep learning-based denoising of low-dose SPECT myocardial perfusion images: quantitative assessment and clinical performance." (in eng), *Eur J Nucl Med Mol Imaging,* Vol. 49 (No. 5), pp. 1508-22, Apr (2022).

92- H. Xie *et al.*, "Deep-Learning-Based Few-Angle Cardiac SPECT Reconstruction Using Transformer." (in eng), *IEEE Trans Radiat Plasma Med Sci,* Vol. 7 (No. 1), pp. 33-40, Jan (2023).

93- Tin Vlašić, Tomislav Matulić, and Damir Seršić, "Estimating uncertainty in pet image reconstruction via deep posterior sampling." *arXiv preprint arXiv:2306.04664,* (2023).

94- Zixin Tang, Caiwen Jiang, Zhiming Cui, and Dinggang Shen, "A New Paradigm for Low-Dose PET/CT Reconstruction with Mamba-Powered Progressive Network and Physics-Informed Consistency." in *International Conference on Medical Image Computing*

and Computer-Assisted Intervention, (2025): *Springer*, pp. 3-12.

95- M. S. Azimi *et al.*, "Deep learning-based partial volume correction in standard and low-dose positron emission tomography-computed tomography imaging." (in eng), *Quant Imaging Med Surg,* Vol. 14 (No. 3), pp. 2146-64, Mar 15 (2024).

96- H. Marquis, K. Erlandsson, and I. Buvat, "Addressing Partial Volume Effects in Clinical PET Quantification: Modern Correction Strategies and Challenges." (in eng), *PET Clin,* Vol. 20 (No. 4), pp. 407-22, Oct (2025).

97- K. Matsubara, M. Ibaraki, T. Kinoshita, and Initiative Alzheimer's Disease Neuroimaging, "DeepPVC: prediction of a partial volume-corrected map for brain positron emission tomography studies via a deep convolutional neural network." (in eng), *EJNMMI Phys,* Vol. 9 (No. 1), p. 50, Jul 30 (2022).

98- M. Azimi *et al.*, "Attention-based deep neural network for partial volume correction in brain (18)F-FDG PET imaging." (in eng), *Phys Med,* Vol. 119p. 103315, Mar (2024).

99- H. Arabi and H. Zaidi, "Deep learning-based metal artefact reduction in PET/CT imaging." (in eng), *Eur Radiol,* Vol. 31 (No. 8), pp. 6384-96, Aug (2021).

100- A. Z. Kyme and R. R. Fulton, "Motion estimation and correction in SPECT, PET and CT." (in eng), *Phys Med Biol,* Vol. 66 (No. 18), Sep 15 (2021).

101- T. Zeng *et al.*, "Supervised Deep Learning for Head Motion Correction in PET." (in eng), *Med Image Comput Comput Assist Interv,* Vol. 13434pp. 194-203, Sep (2022).

102- C. Li, L. A. Polson, X. Wu, Y. Zhang, C. Uribe, and A. Rahmim, "Optical surface information-based respiratory phase-sorting and motion-incorporated reconstruction for SPECT imaging." (in eng), *Med Phys,* Vol. 52 (No. 6), pp. 4330-40, Jun (2025).

103- Y. Lu *et al.*, "Deep learning-aided respiratory motion compensation in PET/CT: addressing motion induced resolution loss, attenuation correction artifacts and PET-CT misalignment." (in eng), *Eur J Nucl Med Mol Imaging,* Vol. 52 (No. 1), pp. 62-73, Dec (2024).

104- T. Li, M. Zhang, W. Qi, E. Asma, and J. Qi, "Motion correction of respiratory-gated PET images using deep learning based image registration framework." (in eng), *Phys Med Biol,* Vol. 65 (No. 15), p. 155003, Jul 30 (2020).

105- A. Artesani, A. Bruno, F. Gelardi, and A. Chiti, "Empowering PET: harnessing deep learning for improved clinical insight." (in eng), *Eur Radiol Exp,* Vol. 8 (No. 1), p. 17, Feb 7 (2024).

106- W. Quanyang *et al.*, "Artificial intelligence in lung cancer screening: Detection, classification, prediction, and prognosis." (in eng), *Cancer Med,* Vol. 13 (No. 7), p. e7140, Apr (2024).

107- H. Arabi *et al.*, "Comparative study of algorithms for synthetic CT generation from MRI: Consequences for MRI-guided radiation planning in the pelvic region." (in eng), *Med Phys,* Vol. 45 (No. 11), pp. 5218-33, Nov (2018).

108- Melika Daraee, Elham Saeedzadeh, Pardis Ghaffarian, and Hossein Arabi, "A Novel Neural Network for Joint Lesion Segmentation and Confidence Score Generation from PET Image." in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, (2022): *IEEE*, pp. 1-3.

109- Hossein Arabi and Habib Zaidi, "Inclusive-Exclusive Model Training Framework to Jointly Perform Semantic Segmentation and Uncertainty Map Estimation." in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, (2022): *IEEE*, pp. 1-3.

110- Hossein Arabi and Habib Zaidi, "Confidence Score Estimation in Machine Learning-based Automated Lesion Segmentation from PET Images." in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, (2022): *IEEE*, pp. 1-3.

111- Masoud Noroozi, Hossein Arabi, and Ali Reza Karimian, "Assessment of Transformer-Based Models for Cardiac Systolic Abnormality Segmentation in Cardiac Catheterization X-Ray Images." in *2024 31st National and 9th International Iranian Conference on Biomedical Engineering (ICBME)*, (2024): *IEEE*, pp. 129-34.

112- M. Champendal, H. Muller, J. O. Prior, and C. S. Dos Reis, "A scoping review of interpretability and explainability concerning artificial intelligence methods in medical imaging." (in eng), *Eur J Radiol,* Vol. 169p. 111159, Dec (2023).

113- A. Salih *et al.*, "Explainable Artificial Intelligence and Cardiac Imaging: Toward More Interpretable Models." (in eng), *Circ Cardiovasc Imaging,* Vol. 16 (No. 4), p. e014519, Apr (2023).

114- S. Usmani *et al.*, "Deep learning (DL)-based advancements in prostate cancer imaging: Artificial intelligence (AI)-based segmentation of (68)Ga-PSMSA PET for tumor volume assessment." (in eng), *Precis Radiat Oncol,* Vol. 9 (No. 2), pp. 120-32, Jun (2025).

115- Aya Farrag, Gad Gad, Zubair Md Fadlullah, Mostafa M Fouda, and Maazen Alsabaan, "An explainable AI system for medical image segmentation with preserved local resolution: Mammogram tumor segmentation." *IEEE Access,* Vol. 11pp. 125543-61, (2023).

116- Dangui Yang, Yetong Wang, Yimeng Ma, and Houqun Yang, "A Multi-Scale Interpretability-Based PET-CT Tumor Segmentation Method." *Mathematics,* Vol. 13 (No. 7), p. 1139, (2025).

117- F. Yang, B. Lei, Z. Zhou, T. A. Song, V. Balaji, and J. Dutta, "AI in SPECT Imaging: Opportunities and Challenges." (in eng), *Semin Nucl Med,* Vol. 55 (No. 3), pp. 294-312, May (2025).

118- S. Toumaj, A. Heidari, and N. Jafari Navimipour, "Leveraging explainable artificial intelligence for transparent and trustworthy cancer detection systems." (in eng), *Artif Intell Med,* Vol. 169p. 103243, Nov (2025).

119- Y. Abas Mohamed, B. Ee Khoo, M. Shahrimie Mohd Asaari, M. Ezane Aziz, and F. Rahiman Ghazali, "Decoding the black box: Explainable AI (XAI) for cancer diagnosis, prognosis, and treatment planning-A state-of-the art systematic review." (in eng), *Int J Med Inform,* Vol. 193p. 105689, Jan (2025).

120- C. Jiang *et al.*, "An explainable transformer model integrating PET and tabular data for histologic grading and prognosis of follicular lymphoma: a multi-institutional digital biopsy study." (in eng), *Eur J Nucl Med Mol Imaging,* Vol. 52 (No. 7), pp. 2384-96, Jun (2025).

121- F. Duan, M. Zhang, C. Yang, X. Wang, and D. Wang, "Non-invasive Prediction of Lymph Node Metastasis in NSCLC Using Clinical, Radiomics, and Deep Learning Features From (18)F-FDG PET/CT Based on Interpretable Machine Learning." (in eng), *Acad Radiol,* Vol. 32 (No. 3), pp. 1645-55, Mar (2025).

122- T. Luo *et al.*, "Improved prognostication of overall survival after radiotherapy in lung cancer patients by an interpretable machine learning model integrating lung and tumor radiomics and clinical parameters." (in eng), *Radiol Med,* Vol. 130 (No. 1), pp. 96-109, Jan (2025).

123- Elias Turner, Penelope Bailey, Roman Carter, Lila Ward, Sawyer Ross, and Aurora Kelly, "Explainable AI in radiation therapy prediction." (2024).

124- S. Cui *et al.*, "Interpretable artificial intelligence in radiology and radiation oncology." *Br J Radiol,* Vol. 96 (No. 1150), p. 20230142, Oct (2023).

125- Zhenyu Yang, "Explainable Artificial Intelligence Techniques in Medical Imaging Analysis." *Duke University*, (2023).

126- Y. Luo, H. H. Tseng, S. Cui, L. Wei, R. K. Ten Haken, and I. El Naqa, "Balancing accuracy and interpretability of machine learning approaches for radiation treatment outcomes modeling." *BJR Open,* Vol. 1 (No. 1), p. 20190021, (2019).

127- A. Hosny *et al.*, "Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study." *PLoS Med,* Vol. 15 (No. 11), p. e1002711, Nov (2018).

128- S. Cui, R. K. Ten Haken, and I. El Naqa, "Integrating Multiomics Information in Deep Learning Architectures for Joint Actuarial Outcome Prediction in Non-Small Cell Lung Cancer Patients After Radiation Therapy." *Int J Radiat Oncol Biol Phys,* Vol. 110 (No. 3), pp. 893-904, Jul 1 (2021).

129- J. Zhang *et al.*, "An Interpretable Planning Bot for Pancreas Stereotactic Body Radiation Therapy." *Int J Radiat Oncol Biol Phys,* Vol. 109 (No. 4), pp. 1076-85, Mar 15 (2021).

130- K. J. Lafata *et al.*, "Association of pre-treatment radiomic features with lung cancer recurrence following stereotactic body radiation therapy." *Phys Med Biol,* Vol. 64 (No. 2), p. 025007, Jan 8 (2019).

131- H. Ji *et al.*, "Post-Radiotherapy PET Image Outcome Prediction by Deep Learning Under Biological Model Guidance: A Feasibility Study of Oropharyngeal Cancer Application." *Front Oncol,* Vol. 12p. 895544, (2022).

132- C. Wang *et al.*, "Dose-Distribution-Driven PET Image-Based Outcome Prediction (DDD-PIOP): A Deep Learning Study for Oropharyngeal Cancer IMRT Application." *Front Oncol,* Vol. 10p. 1592, (2020).

133- A. Pahud de Mortanges *et al.*, "Orchestrating explainable artificial intelligence for multimodal and longitudinal data in medical imaging." *NPJ Digit Med,* Vol. 7 (No. 1), p. 195, Jul 22 (2024).

134- Luís Pinto-Coelho, "How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications." *Bioengineering,* Vol. 10 (No. 12), p. 1435, (2023).

135- Anna Markella Antoniadi *et al.*, "Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review." *Applied Sciences,* Vol. 11 (No. 11), p. 5088, (2021).

136- Tim Räz, Aurélie Pahud De Mortanges, and Mauricio Reyes, "Explainable AI in medicine: challenges of integrating XAI into the future clinical routine." *Frontiers in Radiology,* Vol. 5p. 1627169, (2025).

137- S. N. Saw, Y. Y. Yan, and K. H. Ng, "Current status and future directions of explainable artificial intelligence in medical imaging." *Eur J Radiol,* Vol. 183p. 111884, Feb (2025).

138- J. Abrantes and P. Rouzrokh, "Explaining explainability: The role of XAI in medical imaging." *Eur J Radiol,* Vol. 173p. 111389, Apr (2024).

139- E. Rajabi and K. Etminani, "Knowledge-graph-based explainable AI: A systematic review." *J Inf Sci,* Vol. 50 (No. 4), pp. 1019-29, Aug (2024).

140- Tianjian Guo, Indranil R Bardhan, Ying Ding, and Shichang Zhang, "An explainable artificial intelligence approach using graph learning to predict intensive care unit length of stay." *Information Systems Research,* Vol. 36 (No. 3), pp. 1478-501, (2025).

141- Junlin Hou *et al.*, "Self-explainable ai for medical image analysis: A survey and new outlooks." *arXiv preprint arXiv:2410.02331,* (2024).

142- Laura Bergomi *et al.*, "Which explanations do clinicians prefer? A comparative evaluation of XAI understandability and actionability in predicting the need

for hospitalization." *BMC Medical Informatics and Decision Making,* Vol. 25 (No. 1), pp. 1-13, (2025).

143- Tabea E Röber, Rob Goedhart, and Sİ Birbil, "Clinicians' Voice: Fundamental Considerations for XAI in Healthcare." *arXiv preprint arXiv:2411.04855,* (2024).

144- A. Brankovic *et al.*, "Clinician-informed XAI evaluation checklist with metrics (CLIX-M) for AI-powered clinical decision support systems." *NPJ Digit Med,* Vol. 8 (No. 1), p. 364, Jun 14 (2025).