applied sciences

MDPI

Review

A Review on SPECT Myocardial Perfusion Imaging Attenuation Correction Using Deep Learning

Ioannis D. Apostolopoulos [1], **Nikolaos I. Papandrianos** [1], **Elpiniki I. Papageorgiou** [1,*] and **Dimitris J. Apostolopoulos** [2]

[1] Department of Energy Systems, University of Thessaly, Gaiopolis Campus, 41500 Larisa, Greece; ece7216@ac.upatras.gr (I.D.A.); npapandrianos@uth.gr (N.I.P.)
[2] Laboratory of Nuclear Medicine, University General Hospital of Patras, 26504 Patras, Greece; dimap@med.upatras.gr
* Correspondence: elpinikipapageorgiou@uth.gr

Abstract

Attenuation correction (AC) is an essential process in Single Photon Emission Computed Tomography (SPECT) myocardial perfusion imaging (MPI), an established imaging method for assessing coronary artery disease. Conventional AC approaches typically require CT scans, supplementary hardware, intricate reconstruction, or segmentation processes, which can hinder their clinical applicability. Recently, deep learning (DL) techniques have emerged as alternatives, allowing for the direct learning of attenuation patterns from non-AC (NAC) imaging data. This review explores the existing literature on DL-based AC methods for SPECT MPI. We highlight high-performing models, including attention-gated U-Net conditional Generative Adversarial Networks (GANs), and evaluate their validation methods. Although significant advancements have been achieved, numerous challenges persist, which are thoroughly discussed.

Keywords: attenuation correction; myocardial perfusion imaging; deep learning

check for updates

Academic Editors: Julio Garcia Flores and Pedro Couto

Received: 20 September 2025
Revised: 13 October 2025
Accepted: 20 October 2025
Published: 21 October 2025

Citation: Apostolopoulos, I.D.; Papandrianos, N.I.; Papageorgiou, E.I.; Apostolopoulos, D.J. A Review on SPECT Myocardial Perfusion Imaging Attenuation Correction Using Deep Learning. *Appl. Sci.* **2025**, *15*, 11287. https://doi.org/10.3390/app152011287

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

SPECT MPI is an established, fundamental imaging modality for the diagnosis and management of coronary artery disease (CAD) [1,2]. However, its diagnostic accuracy can be undermined by inherent limitations, such as photon attenuation [3]. Photon attenuation arises from the absorption and scattering of emitted gamma photons by tissues of varying densities, leading to artifacts that obscure the true distribution of radiotracers in the myocardium [3]. This phenomenon affects image quality and complicates the quantitative assessment of myocardial blood flow [4,5]. As a result, physicians may misinterpret perfusion defects.

Since the introduction of hybrid systems in the late 1990s, which incorporate a computerized tomography (CT) scanner into a conventional gamma-camera, CT is now used for AC of SPECT slices [6]. From its initial application up to today, single, 4, 6, 16 and even 64 slice CT scanners have been coupled with gamma-cameras in different commercially available SPECT/CT machines. CT produces images that serve as a transmission map, which is then applied to SPECT data to correct photon attenuation [7]. Although the attenuation pattern of gamma photons by various tissues is not identical to that of X-rays, this AC method has gained wide acceptance, due to the fast CT acquisition procedure, and other advantages provided by CT images, such as the exact localization of SPECT findings in space and the additional diagnostic information of anatomical imaging [8]. However, CT artifacts (from implanted metallic devices, respiratory motion and other sources) and misregistration of sequentially acquired CT and SPECT images can cause AC inaccuracies [9,10].

Conventional attenuation correction in myocardial perfusion SPECT has relied on transmission information from CT, on segmentation or model-based approaches, and on uniform attenuation maps. These techniques have improved image quality but they introduce practical burdens that limit adoption in routine workflows [11]. Additional radiation and cost from CT, sensitivity to misregistration between emission and transmission studies, and the need for specialized hardware remain important constraints. Segmentation and model-based methods often depend on assumptions that do not hold across diverse body types and camera configurations. These constraints motivate alternative strategies that are robust to patient variability and that minimize additional acquisition and hardware requirements [12].

Deep Learning (DL) [13] approaches have influenced the field of medical imaging, offering new solutions to longstanding challenges in SPECT MPI AC [14–23]. Convolutional neural networks (CNNs) [24] and other deep architectures have demonstrated remarkable capabilities in learning complex, non-linear mappings from raw imaging data to corrected outputs [25]. However, the integration of deep learning into SPECT MPI AC has only recently been investigated. Recent developments provide a data driven route to AC that learns attenuation patterns directly from non-corrected emission data. Such approaches can reduce reliance on supplementary transmission scans, shorten workflow, and can be trained to preserve clinically meaningful perfusion patterns. Early studies report performance that approaches CT-based correction on diagnostic metrics while avoiding additional acquisitions. These factors make DL particularly relevant now that large clinical archives and modern camera systems are widely available.

This review paper aims to provide a comprehensive overview of the state-of-the-art in DL-based AC for SPECT MPI.

Research Questions

The review addresses the following questions based on the literature:

- Which DL architectures and algorithms have demonstrated the highest performance in AC?
- To what extent do studies integrate non-image data (e.g., patient demographics, clinical parameters) into DL models, and how does this inclusion influence model accuracy and generalizability?
- What quantitative metrics are employed to assess model performance and error?
- Do researchers utilize established MPI quantification metrics, such as Summed Stress Score (SSS), Summed Difference Score (SDS), Summed Rest Score (SRS), or Total Perfusion Deficit (TPD) to evaluate the quality of DL-generated AC images?
- What are the typical sizes and compositions of training and validation datasets reported in the literature?
- How frequently do studies employ independent external populations for additional testing?

We posed these questions to better understand how DL is currently used for AC. There is a lot of variation in the models and techniques being applied in related tasks [26], so identifying which architectures perform best can help steer future work in a more focused direction.

We also argue that how often non-image data is used as inputs to the models may be critical, since it could potentially have a substantial impact on the AC accuracy of the models. Another important matter is how researchers measure the quality and precision of the generated AC images, because of the lack of a standardized set of performance metrics in the field [27]. Similarly, understanding whether established MPI quantification metrics are used provides insight into how clinically meaningful the results are.

Finally, reviewing the utilized datasets sizes, diversity, and the use of external validation reveals a lot about how trustworthy and transferable these models are.

2. Background

2.1. Single Photon Emission Computed Tomography

SPECT is a nuclear imaging technique used in clinical practice for functional imaging, particularly in cardiovascular diagnostics [10]. It enables the visualization and quantification of physiological processes by detecting gamma rays emitted from radiopharmaceuticals administered to a patient. In MPI [28], SPECT holds a major role in evaluating CAD, helping clinicians detect ischemia, assess myocardial viability, and stratify patient risk [3].

SPECT imaging involves three essential steps: radiotracer injection, image acquisition using a gamma camera, and computational image reconstruction. Technetium-99m (^{99m}Tc) labeled compounds such as ^{99m}Tc-sestamibi or ^{99m}Tc-tetrofosmin, and Thallium-201 (Tl-201), which distribute in the myocardium proportionally to blood flow are the most used radiopharmaceuticals [29]. These isotopes emit gamma photons that are then detected externally to form a projection dataset, which is subsequently reconstructed into a three-dimensional image of the radiotracer distribution within the heart [30].

Compared to Positron Emission Tomography (PET), SPECT is more cost-effective and widely available, though it typically offers lower spatial and temporal resolution [31]. Nevertheless, its clinical utility is considerable due to its diagnostic and prognostic value, particularly in resource-constrained settings [32]. The technique is especially powerful when combined with stress testing, revealing regional differences in myocardial perfusion that can indicate significant coronary stenosis.

A major limitation of SPECT is the physical phenomenon of photon attenuation [33]. As gamma rays travel through the body, they may be absorbed or scattered by tissues of varying densities (e.g., the diaphragm, breasts, ribs), leading to non-uniformities in the detected signal [34]. These effects often create false perfusion defects or obscure true abnormalities and, as a result, undermine diagnostic accuracy. This limitation necessitates AC, which compensates for photon loss and enhances the fidelity of reconstructed images.

Historically, AC in SPECT has relied on the incorporation of anatomical imaging modalities such as CT [35]. Hybrid SPECT/CT systems were introduced in the late 1990s and gained clinical popularity due to their ability to generate attenuation maps from X-ray data [6]. These maps are subsequently used during image reconstruction to apply spatially varying correction factors. CT-based AC has improved diagnostic accuracy by reducing artifacts and enabling anatomical correlation, but it introduced several practical concerns, including additional radiation exposure, potential image misregistration due to respiratory or cardiac motion, and logistical challenges related to hardware cost and maintenance.

Efforts have been made to explore non-CT-based AC strategies, such as segmentation-based methods, transmission scans using radioactive sources (e.g., gadolinium-153), and statistical modeling approaches [36]. These methods come with drawbacks, such as reliance on anatomical assumptions, limited generalizability, and increased scan duration. More recently, DL techniques have shown promise in learning complex attenuation patterns directly from emission data [33], offering a new direction for SPECT AC that may bypass the need for additional hardware.

Another SPECT practice, recently introduced, is the camera and collimator technology [37]. Most conventional systems use parallel-hole collimators, but newer dedicated cardiac systems, such as those based on cadmium–zinc–telluride (CZT) detectors, utilize multi-pinhole or IQ·SPECT collimators with multifocal geometries. These advancements provide improved spatial resolution and sensitivity, though they also present unique challenges for AC due to altered photon trajectories and sensitivity profiles.

Beyond replacing CT in AC, AI can improve CT inputs. Deep denoising allows lower-dose CT while preserving HU fidelity. Supervised and adversarial metal-artifact reduction reduce streaks that bias μ and learned deformable registration can reconcile respiratory phase differences between CT and emission data. HU-harmonization models further stabilize bilinear scaling across scanners and kVp choices, yielding μ-maps that are more consistent for downstream reconstruction. These supportive AI components are modality-agnostic and particularly helpful when CT quality or registration is the dominant error source in SPECT/CT workflows.

Furthermore, image reconstruction algorithms have evolved considerably. Traditional filtered back-projection (FBP) has largely been supplanted by iterative reconstruction methods such as ordered-subset expectation maximization (OSEM) [38], which can incorporate physical modeling of detector response, scatter, and attenuation [39]. These reconstructions can be further enhanced by integrating DL-generated attenuation maps or direct emission-to-corrected image mappings. Choice of radiopharmaceutical influences photon energy and scatter fraction, thereby affecting attenuation magnitude and the stability of AC across stress/rest and tracer protocols. Collimator geometry (e.g., LEHR, multipinhole, IQ·SPECT) alters path-length distributions and sensitivity profiles, which can exacerbate regional attenuation (e.g., inferior wall adjacent to the diaphragm) and complicate μ-map generation. Reconstruction methods (FBP vs. OSEM with resolution recovery and scatter modeling) determine whether AC is applied as a physics model during iteration or post hoc in image space; model-based reconstruction with accurate attenuation terms generally yields better uniformity but is sensitive to μ-map errors and misregistration.

From a clinical standpoint, SPECT MPI continues to be indispensable in cardiology, supported by its inclusion in numerous international guidelines and appropriate criteria. It provides semi-quantitative measures such as the SSS, SRS, and TPD, which aid in diagnosis and therapeutic planning. Incorporating reliable AC, whether via CT or emerging DL methods, is essential to preserving the modality's clinical impact and enabling quantitative assessments that correlate with invasive angiographic findings.

PET commonly leverages transmission scans or CT with well-characterized linear attenuation at 511 keV, stable time-of-flight statistics, and routine motion management. These features historically informed SPECT AC practices. Some studies have proposed AI methods for AC in PET scans [40–44] and brain perfusion SPECT [45,46]. SPECT's lower photon energy and modality-specific collimation lead to different scatter/attenuation behaviors, so SPECT requires tailored AC solutions, particularly for myocardial perfusion where soft-tissue attenuation is non-uniform and motion is frequent.

2.2. Deep Learning for Attenuation Correction

DL, particularly in the form of CNNs [47,48], has transformed image analysis and synthesis by enabling data-driven, end-to-end learning of complex relationships within high-dimensional data. In medical imaging, DL captures subtle spatial patterns, corrects for noise and artifacts, and models highly non-linear transformations [49,50]. These capabilities make it particularly suitable for addressing the challenge of AC in SPECT MPI (Figure 1) and other SPECT MPI related tasks, such as CAD diagnosis [51–54].

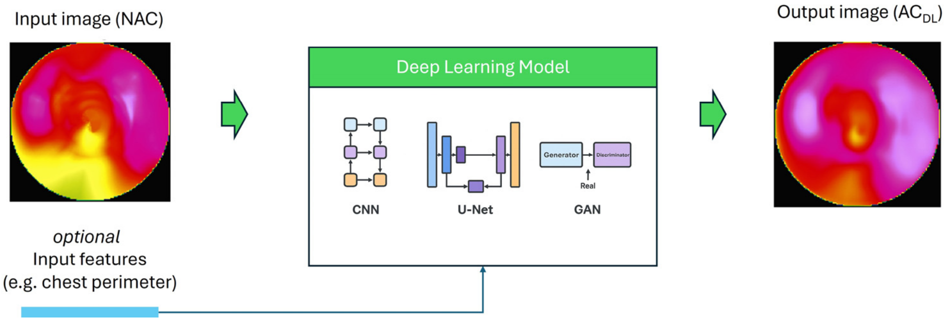

Figure 1. General schematic of DL-based AC in MPI. Inputs include non corrected short axis volumes or polar maps and may optionally include auxiliary channels such as scatter windows or paired stress and rest studies. A generator network such as a U-Net with attention gates produces either directly corrected emissions or a patient specific attenuation map. When a discriminator is used it enforces anatomical and textural realism. The outputs are used either as corrected images for quantification or as attenuation factors within iterative reconstruction. Evaluation uses voxel level similarity, polar map agreement, and clinical indices.

Deep networks support non-linear mappings that counter non uniform soft tissue attenuation and they capture multi scale spatial context that reflects diaphragm position and breast tissue distribution. In SPECT AC this translates into recovery of inferior and lateral wall activity while preserving true defects. Conditional adversarial training and attention gated U-Net generators focus capacity on myocardium and attenuation related structures and have shown improved agreement with CT-based correction and improved diagnostic performance in external cohorts.

DL does not rely on explicit modeling of photon propagation or rigid anatomical priors. Instead, it learns the underlying attenuation behavior directly from data [46]. This is achieved through a hierarchical representation of features, allowing the network to build abstract concepts from raw input, such as identifying spatial patterns that indicate tissue density or typical anatomical configurations. However, true AC images are required for DL to learn these transformations from NAC to AC.

A key advantage of DL in this context is that it models non-linear relationships between the NAC input and the true AC image. Deep neural networks approximate any continuous function given sufficient data and appropriate architecture [55]. The latter is what makes them uniquely positioned to correct for non-uniform attenuation patterns across diverse patient anatomies and imaging conditions.

Moreover, convolutional architectures like U-Nets and ResNets [56,57] are designed to capture spatial dependencies at multiple scales. In the context of AC, such network topologies consider not only local voxel/pixel intensities but also larger anatomical context (e.g., diaphragm position, breast tissue distribution, or lateral wall thickness) to infer how photon attenuation might have distorted the measured signal. This context-awareness is critical for accurately reconstructing perfusion patterns that are not visible in raw NAC data alone. cGANs further constrain texture/anatomy realism via an adversarial discriminator. Indirect pipelines that predict μ-maps before iterative reconstruction provide physics-consistent AC and facilitate quality control of the μ-map. Direct NAC to AC image translation offers faster inference and simpler integration but needs careful bias control.

DL is also highly amenable to volumetric and multi-channel inputs, enabling the integration of different imaging modalities (e.g., scatter windows, stress/rest pairs), anatomical priors, or even demographic data into a unified predictive framework. Using deep networks enables end-to-end training, where the objective function can be tailored to clinical goals. For example, training loss can be defined in terms of voxel-wise error, structural similarity, or even clinical decision metrics like perfusion scores.

Despite the positives, DL models are inherently dependent on large volumes of data. Access to well-annotated, high-quality paired NAC and AC datasets is limited due to privacy concerns, institutional variability, and the need for CT-based ground truth. Additionally, DL models may be prone to overfitting if trained on narrowly sampled populations or scanner types, potentially limiting generalization. Other issues DL faces include interpretability [58], reproducibility [59], and computational expenses.

AC directly alters counts and polar-map scores. Hence, interpretability cannot be optional. Simple, task-linked tools, such as voxel-level "change maps" that audit how much correction was applied, saliency overlays highlighting attenuation-driven regions, and calibrated uncertainty maps, might help readers judge when to trust DL-AC and when to defer to CT-AC. Reproducibility likewise requires deterministic or well-documented training/inference pipelines (including seeds and software versions), pre-specified evaluation protocols, release of code/model weights when possible, and reporting that spans voxel-level similarity, segment-level bias, and patient-level diagnostic metrics.

Current evidence supports DL AC as a complement to established pipelines, with pathway to a replacement contingent on multi-center validation, standardized clinical metrics, and prospective workflow studies.

3. Methodology of Literature Review

3.1. Search Strategy

A structured literature review was conducted using DL approaches to identify relevant studies on SPECT AC. A comprehensive search was performed across multiple scientific databases, including PubMed, Scopus, and Web of Science, using a combination of keywords such as "SPECT," "myocardial perfusion," "attenuation correction," "deep learning," and "artificial intelligence".

An identical generic search string was applied across databases as follows: (("single photon emission computed tomography" OR "single-photon emission computed tomography" OR SPECT) AND ("myocardial perfusion" OR MPI) AND ("attenuation correction" OR attenuation OR "SPECT/CT") AND ("deep learning" OR "machine learning" OR "artificial intelligence" OR")).

Searches were last run on 20 February 2025. Records from all databases were exported on the same day and deduplicated prior to screening.

3.2. Inclusion and Exclusion Criteria

A set of predefined inclusion and exclusion criteria were established and consistently applied during the selection process to ensure the relevance and quality of the literature reviewed. Studies were eligible for inclusion in this review if they satisfied all of the following conditions:

- The primary objective of the study was AC in SPECT MPI.
- The methodology involved the use of DL.
- The article was peer-reviewed and published in the English language.
- The publication date was after 2020.

Studies were excluded from the review if they met any of the following conditions:

- The publication was a review article, editorial, commentary, or a conference abstract without an accompanying full-text manuscript.
- The study was a duplicate of another included work.

3.3. Paper Selection

The selection of relevant studies was conducted through a systematic literature search. The procedure involved several sequential steps. Firstly, we removed all studies before 2020. All the remaining retrieved records were imported into Zotero. Duplicate entries were identified and removed to ensure each study was evaluated only once. Two independent reviewers (I.D.A. and N.P.) screened the titles and abstracts of the remaining studies to identify potentially relevant articles. Discrepancies between reviewers were resolved through discussion or consultation with a third reviewer (D.J.A.). The full texts of the shortlisted articles were then assessed against the inclusion and exclusion criteria described in Section 2.2. Studies that did not meet all inclusion criteria were excluded, with reasons documented.

The final set of articles included in the review met all eligibility criteria. A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram was prepared to summarize the selection process visually (Figure 2). Sixteen studies qualified for this review based on the selection process. Two reviewers (I.D.A. and E.P.) screened the final full texts and discrepancies were resolved by discussion or adjudication by a third reviewer (D.J.A.).

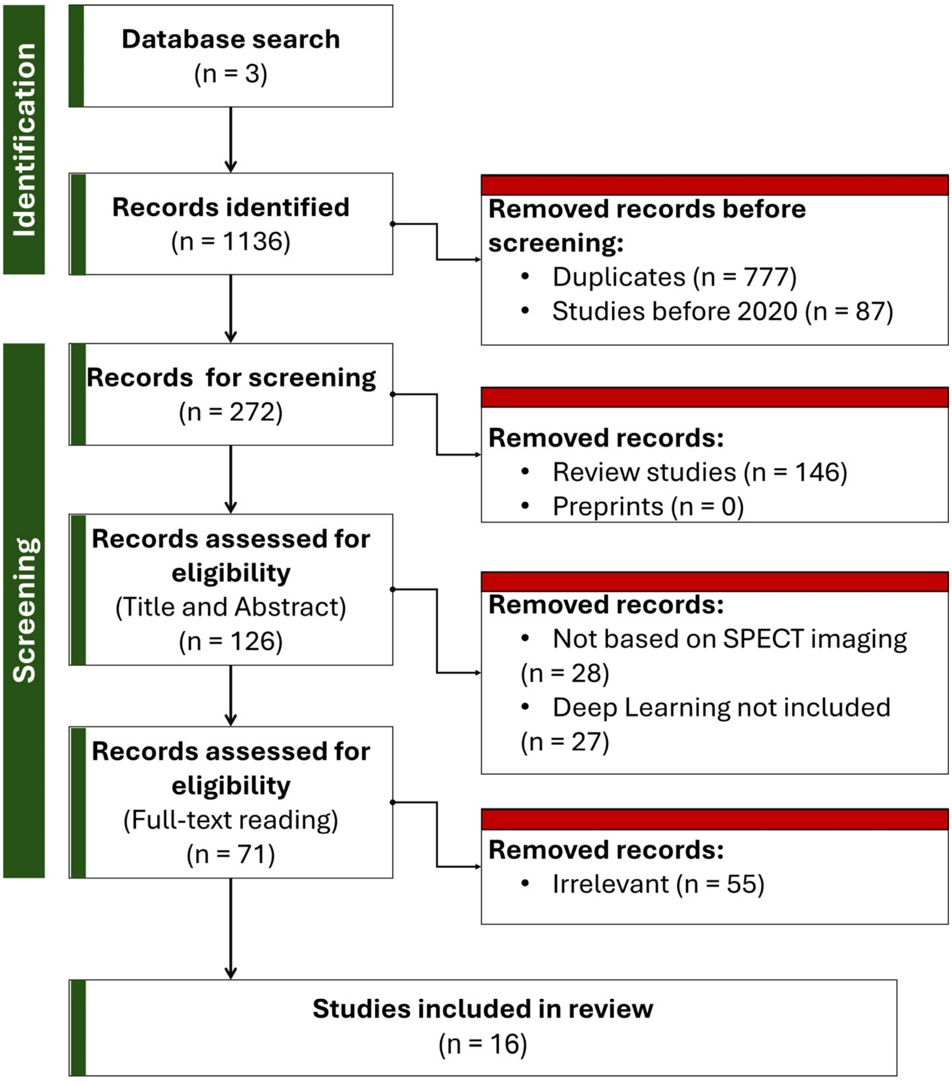

Figure 2. PRISMA flow diagram.

4. Review Findings

In this section, we provide a short summary of each eligible study.

Hagio et al. [60] developed and validated a DL algorithm to perform "virtual" attenuation correction. Using a dataset of over 11,000 stress SPECT studies with paired CT-based attenuation-corrected (CTAC) and NAC images, they trained a modified U-Net CNN to generate attenuation-corrected polar maps directly from NAC inputs.

In their findings, the DLAC polar maps showed markedly improved correlation with CTAC maps compared to NAC alone ($R^2 = 0.85$ vs. 0.68, $p < 0.001$). When assessed against invasive coronary angiography (ICA) as the diagnostic gold standard, DLAC demonstrated a significantly higher area under the ROC curve (AUC = 0.827) than NAC (AUC = 0.780, $p = 0.012$), and performed comparably to CTAC (AUC = 0.851, $p = 0.208$ vs. DLAC). Specificity and accuracy were also significantly improved with DLAC, especially at high sensitivity thresholds (e.g., 88% sensitivity yielded 18.9% specificity improvement over NAC). A key strength of this study is its large training population, which enables robust model development and evaluation. Moreover, clinical utility was assessed using meaningful diagnostic metrics such as TPD and comparison to angiographic outcomes, rather than relying solely on image similarity. Visual assessments confirmed that DLAC preserved true perfusion defects while correcting soft-tissue artifacts. However, the work is limited by its single-center design and reliance on data acquired using a single camera system, which may affect generalizability. The model did not incorporate non-image patient-specific data (e.g., BMI, clinical history), and evaluation was confined to TPD without broader myocardial scoring indices such as SRS or SDS.

Prieto Canalejo et al. [61] presented a DL-based method to generate synthetic attenuation correction maps (ACMs) directly from NAC SPECT emission images. The authors retrospectively collected 384 myocardial perfusion SPECT/CT studies using 99mTc-sestamibi: 312 for training, 64 for validation, and 66 held out for testing. A 2D U-Net with five encoder–decoder layers was trained to predict linear attenuation coefficient maps at 140 keV from three-channel NAC inputs (adjacent slices as channels) using a composite loss (structural similarity + mean relative error). On the independent test set, the synthetic ACMs achieved Mean Structural Similarity Index (SSIM) MSSIM 0.97 ± 0.01 and Normalized Mean Absolute Error (NMAE) 3.08 ± 1.26%, while reconstructed emission images corrected with these maps reached MSSIM 0.99 ± 0.003 and NMAE 0.23 ± 0.13%. Bland–Altman analysis showed voxel-level limits of agreement within [–9.04, 9.00]% for emission data and segment-level agreement [–11, 10]% for 17-segment polar maps. Semi-quantitative SSS comparisons yielded 95% limits [–2.8, 3.0] with only 3% of patients shifting perfusion categories. The model attained performance comparable to prior larger studies using far fewer training examples, suggesting efficient learning of attenuation patterns.

Huxohl and colleagues [62] retrospectively collected 150 cardiac SPECT studies from 103 patients acquired on a Siemens Symbia Intevo IQ·SPECT system and trained a 3D U-Net within a cGAN framework to predict CT-based attenuation maps from non-attenuation-corrected reconstructions. They optimized hyperparameters via random search (215 trials), found that an L1 distance loss and a classification-network discriminator yielded the best NMAE (0.020 ± 0.007 on test data), and demonstrated high agreement between DL AC and CTAC polar maps (segment-wise APE 1.155 ± 0.769, R = 0.97). A strength of this work is its focus on a clinically prevalent IQ·SPECT collimator, whose multifocal geometry and body-covering field of view pose unique challenges, thus extending prior methods beyond pinhole or LEHR systems. The indirect approach (map prediction rather than direct AC reconstruction) adds interpretability and fault tolerance, since errors in fine structures of the map have limited impact on the final reconstruction. Moreover, the extensive random search provides empirical guidance on loss functions and discriminator choice, which can inform future studies. The absence of standard clinical scoring metrics (SSS, SDS, SRS, TPD) leaves uncertain whether the close image-based agreement translates into equivalent diagnostic performance.

Mostafapour et al. [63] retrospectively gathered 99 99mTc-sestamibi MPI-SPECT studies acquired on a GE Discovery NM/CT 670 dual-head SPECT/CT, applied scatter correction, and reconstructed 64 × 64 × 40 ungated images. They trained two 2D deep networks (a dilated-ResNet and a symmetric U-Net) in the NiftyNet framework to directly predict CT-based AC images from NAC inputs, reserving 19 subjects for validation. Quantitatively, both models vastly outperformed the uniform Chang method: U-Net achieved Mean Error (ME) = −4.41 ± 11.85 counts, MAE = 13.65 ± 11.23, SSIM= 0.98 ± 0.05, and TPD = 12.57 ± 8.93% versus CTAC's 12.84 ± 8.63% (Chang SSIM = 0.93 ± 0.09; TPD = 16.68 ± 11.24%). Visual assessments, bias maps, and line profiles confirmed recovery of myocardial uptake, though occasional UNet outliers appeared in low-count cases. A key strength is the inclusion of both low-level image metrics and clinically meaningful perfusion indices (SSS, TPD, defect size).

Shanbhag et al. [64] trained a cGAN with an attention-gated 3D U-Net generator to translate NAC SPECT short-axis volumes into simulated AC images (called DeepAC), using 4886 paired CT-AC studies for training. They then validated DeepAC on 604 patients from two external centers, including 280 with same-day coronary CTA and 324 low-likelihood cases. DeepAC quantitative perfusion (stress TPD) achieved an AUC of 0.79 (95% CI 0.72–0.85) for obstructive CAD, outperforming NAC (AUC 0.70, $p < 0.001$) and matching CTAC (AUC 0.81, $p = 0.196$). The normalcy rate rose to 70.4% with DeepAC versus 54.6% non-AC ($p < 0.001$), and voxel-wise "change analysis" showed substantially less positive count change (artifact correction) needed by DeepAC (median 2.4 counts) than NAC (9.4 counts), closely aligning with CTAC (9.4 counts). A major strength of this work is the use of a large, truly external test set drawn from two independent centers and including both obstructive CAD and defined low-likelihood cohorts, which demonstrates the model's generalizability beyond its training site. Moreover, the voxel-wise change analysis provides an interpretable measure of how DeepAC corrects attenuation artifacts. The model was trained and tested solely on General Electric (GE) solid-state SPECT systems. Hence, its performance on other vendors' cameras or collimator designs is unknown.

Shi et al. [65] developed a 3D cGAN that fuses primary- and scatter-window SPECT reconstructions to predict CT-equivalent attenuation maps. On 25 held-out subjects, their GAN-PS model yielded a global NMAE of 3.60 ± 0.85% against true CT maps, and when used for attenuation correction, produced SPECT images with <1.5% myocardial bias and <1.1% blood-pool bias—differences not statistically significant versus CT-based correction. A key advantage of this work is its innovative use of multi-energy windows to leverage scatter information, coupled with adversarial training to sharpen map fidelity. However, the study is limited by its small, single-center cohort and lack of clinical outcome metrics (e.g., TPD or SSS). Additionally, reliance on full-FOV SPECT volumes may hinder applicability on scanners with truncated views or on protocols lacking dedicated scatter data.

Lim et al. [66] demonstrated the clinical feasibility of a U-Net-based DL model to perform CT-level AC on Tl-201 myocardial perfusion SPECT without additional hardware or radiation. Training on nearly a thousand stress/rest studies, the model translated NAC images into synthetic, attenuation-corrected images with sub-percent MAE, near-unity SSIM (≃0.99), and excellent Peak Signal to Noise Ratio (PSNR) (≈34 dB). On segmental polar-map analysis, the generated images fully restored inferior-wall counts to match CTAC references and eliminated soft-tissue artifacts. In a focused observer study of diaphragmatic attenuation, 95% of the synthetic images were scored as "similar" or "indistinguishable" from CTAC studies. Strengths of this work include its use of clinically routine Tl-201 protocols (where attenuation is worst), rigorous stress/rest training and testing, and both quantitative and blinded visual validation. Limitations include its single-center, retrospective design without external validation, lack of breast-attenuation cases to fully assess women, and as yet unexplored performance in obese patients or on different camera models—areas that warrant further prospective, multicenter study.

Yang et al. [67] introduced FA-ACNet, a novel 3D U-Net-based framework for CT-free AC of cardiac SPECT that aligns multi-scale features extracted from NAC SPECT and low-dose CT images to generate AC images comparable to CTAC references. They demonstrate excellent quantitative performance (Mean Square Error (MSE) of 11.98×10^{-6}, SSIM of 0.9976, and PSNR of 45.54) in an independent test set and strong agreement in summed perfusion scores (SSS/SRS bias ≤ 0.2, 95% LoA within $\pm$ 2.8), while providing visual case examples and Bland–Altman analyses. However, the study's single-center design, reliance on one tracer (^{99m}Tc-sestamibi) and one scanner model, and absence of an external validation cohort restrict generalizability.

Chen et al. [68] systematically compared direct image-to-image and indirect μ-map-based DL strategies for CT-free AC in both dedicated cardiac and general purpose SPECT, employing U-Net as a baseline and an advanced DuRDN architecture to process photopeak and multi-window scatter inputs. They demonstrated that indirect approaches (predicting truncated or full field-of-view μ-maps prior to iterative reconstruction) achieve significantly lower normalized MSE (NMSE) (1.20 ± 0.72% vs. 2.21 ± 1.17% for dedicated SPECT) and polar-map APE (3.24 ± 2.79% vs. 4.77 ± 3.96%) than direct methods, while DuRDN consistently outperformed U-Net in artifact suppression and quantitative fidelity. The study used tailored pipelines that recover out-of-FOV attenuation information for pinhole systems and rigorous multi-metric evaluation, including voxel, segment, and defect-size analyses.

Yang et al. [69] presented a direct DL method for CT-less AC in dedicated cardiac CZT SPECT, training a 3D convolutional encoder–decoder network to map non-corrected SPECT volumes directly to CT-based AC reference images. In 10-fold cross-validation of 100 clinical studies, their DL outputs achieved a marked improvement over uncorrected images: voxel-wise R^2 increased from 0.81 to 0.91, average segmental error from –6.1% to +0.5%, and mean absolute segmental error halved to 3.3% (all $p < 0.001$). A key strength of this work is its demonstration that a single DL step in image space can substantially reduce soft-tissue attenuation artifacts, enabling rapid (∼0.5 s) correction on stand-alone CZT scanners without CT, line sources, or transmission scans, and potentially lowering patient dose and streamlining clinical workflow. The proposed model sometimes over- or under-corrected in atypical uptake patterns (as seen in worst-case examples), reflecting reliance on training distribution and lack of explicit physics modeling.

Torkaman et al. [70] proposed SPECTGAN, a 3D conditional GAN that directly translates NAC cardiac SPECT volumes into AC. In 5-fold cross-validation of 100 dedicated CZT SPECT/CT studies, SPECTGAN reduced normalized Root Mean Squared Error (RMSE) from 0.226 to 0.141 (37.5% error reduction), boosted PSNR by 14.5%, and improved SSIM to 0.995 compared to 0.988 in uncorrected images.

In another study by Torkaman et al. [71], the authors addressed patient-to-patient variability in direct SPECT attenuation correction by evaluating two DL approaches (ResUNet and a Wasserstein CycleGAN) on 100 paired non-corrected/corrected scans via leave-one-out cross-validation. Quantitative metrics (NRMSE, PSNR, SSIM) and visual overlays showed the CycleGAN reduced global error by 8.9% and preserved background anatomy without masking. To further stabilize performance, the authors clustered polar-map features using t-SNE and BIRCH, then trained models on these homogeneous subsets; this yielded up to a 36% reduction in NRMSE variability and a 6.4% overall NRMSE gain when using data most similar to each test case.

Hagio et al. [72] validated an AC CNN on 722 SPECT MPI studies from 74 international sites and three scanner vendors in the phase 3 flurpiridaz-301 trial. They generated polar maps from NAC, then quantified TPD and compared diagnostic performance against invasive coronary angiography. The proposed method achieved a per-patient AUC of 0.752 versus 0.717 for NAC ($p = 0.016$), improving specificity by 6.2% and accuracy by 4.3% without sacrificing sensitivity; results matched expert-reader SSS (AUC = 0.743, $p = 0.913$). This multi-center validation demonstrates DLAC's robustness across diverse hardware and protocols. Limitations of this study included the lack of CT-based ground-truth attenuation maps and the absence of non-image patient features (e.g., BMI subgroup analysis).

Chen and colleagues [73] evaluated DL estimation of attenuation maps for cardiac SPECT/CT using a 3D U-Net fed with both photopeak and one of three scatter windows (4%, 10%, or 30% width) and tested robustness at full-count versus quarter-count levels. The work used 1517 patient studies (905 cases for training, 226 for validation, and 386 for testing). Across all scatter configurations, DL-estimated maps agreed with CT-derived ground truth within 2.1% NMSE (mean bias ≈ 0), with SSIM $\approx$ 0.94–0.99 and PSNR $\approx$ 29–30 dB. Wider scatter windows yielded marginally better map fidelity. When these DL maps were used for AC SPECT reconstruction, slice NMSE fell from ~9% (no AC) to <1%, and polar-map segmental uptake errors were <5%. Performance at 1/4-count was only slightly degraded (NMSE map $\approx$ 0.3%, AC slices $\approx$ 0.05%). This study demonstrates that U-Net-based DL can reliably infer patient-specific attenuation maps across varied scatter sampling schemes and even from reduced-count data, obviating CT acquisition without compromising quantitative accuracy.

Chen et al. [74] presented a CT-free AC method for cardiac SPECT utilizing a customized 3D Dual Squeeze-and-Excitation Residual Dense Network (DuRDN), which innovatively integrates non-image clinical features such as BMI, gender, and scatter window images to enhance the model's performance. A notable strength of this study is the explicit incorporation of patient-specific variables, reflecting an advanced effort toward personalized imaging, and achieving a voxel-wise NMSE of 2.01 ± 1.01%, outperforming a conventional U-Net (2.23 ± 1.20%). Moreover, visual comparisons between predicted and ground-truth AC images confirmed qualitative consistency.

Ochoa-Figueroa et al. [75] assessed the clinical utility of a commercially developed DL-based attenuation correction software (DLACS) for MPI using a CZT dedicated camera. A major strength of the study is the use of ICA as the gold standard for validation, providing clinically meaningful endpoints rather than solely imaging metrics. The application of DLACS improved the diagnostic specificity dramatically from 57% to 86%, while maintaining high sensitivity (91%), resulting in an overall diagnostic accuracy increase from 87% to 90%. These results underscore the potential of DL to reduce false positives and artifact-driven misinterpretations in MPI.

5. Comparative Review

This section provides answers to the research questions of the present study by summarizing the findings of the qualified related work. Table 1 summarizes the characteristics of each qualified study and Table 2 summarizes key findings.

5.1. Which DL Architecture and Algorithm Has Demonstrated the Highest Performance in AC?

According to the reviewed studies, three-dimensional conditional generative adversarial networks (3D cGANs), particularly those featuring attention-gated U-Net generators, have demonstrated the highest AC performance. Attention-gated mechanisms are components within neural network architectures that enable the model to focus selectively on the most informative regions of the input data while suppressing less relevant background features. For the particular 3D cGAN topology for AC, attention gates embedded in the U-Net generator highlight myocardial regions and attenuation-related structures that are most critical for accurate reconstruction. This helps the network allocate its capacity more effectively, reducing noise and artifacts while preserving clinically important perfusion patterns.

A 2023 study [64] employing a 3D attention-gated U-Net within a cGAN framework achieved outstanding results, validated on two independent external cohorts totaling 604 patients, using diagnostic accuracy metrics such as AUC based on TPD. Another high-performing model is the FA-ACNet, developed in 2025 [67], which is based on a dual-encoder 3D U-Net architecture and reported strong internal performance metrics such as SSIM, MSE, and PSNR. However, FA-ACNet lacked external validation. Overall, 3D cGAN architectures and U-Nets, particularly those tested on external populations, appear to demonstrate the most robust performance.

Across the reviewed literature, most direct image-to-image approaches optimize voxelwise L1/MAE (sometimes combined with SSIM), whereas cGAN variants add an adversarial term to promote perceptual realism. Explicit ablations were uncommon. A notable exception is Huxohl et al. [62], who screened losses and discriminator designs and reported their best results with L1 plus a classification discriminator (lowest NMAE on held-out test data). While adversarial training can sharpen textures, the clinical endpoints (e.g., TPD-based AUC or reader-oriented normalcy) are driven more by data diversity and validation design than by a specific loss choice in isolation. Full analytic comparison of loss landscapes and their bias–variance trade-offs are beyond the present scope, but we highlight that robust performance has been achieved with relatively simple L1-dominant objectives.

5.2. To What Extent Do Studies Integrate Non-Image Data (e.g., Patient Demographics, Clinical Parameters) into DL Models, and How Does This Inclusion Influence Model Accuracy and Generalizability?

The integration of non-image data into DL models for AC is currently rare. Most studies strictly utilize image-based inputs without incorporating patient demographics or clinical parameters. Notable exceptions include a 2022 study employing a DuRDN which incorporated features such as BMI, gender, and scatter window images [74]. Another study utilized photopeak and scatter imaging windows as additional inputs, though demographic information was absent [68]. Despite these efforts, there is little evidence that including non-image features significantly enhanced model performance or generalizability. The models demonstrating the strongest external validation results did not use any non-imaging data, suggesting that current advances in model architecture and training design may outweigh the marginal benefits of feature augmentation with patient-specific information.

5.3. What Quantitative Metrics Are Employed to Assess Model Performance and Error?

The reviewed studies employed many quantitative metrics to assess model performance and error. Common voxel-based metrics included the SSIM, NMSE, MAE, NMAE, PSNR, and Pearson correlation coefficients. Several studies also utilized Bland–Altman plots and segment-wise percentage errors to evaluate performance at both voxel and anatomical segment levels. For clinical relevance, receiver operating characteristic (ROC) analysis, area under the ROC curve (AUC), sensitivity, specificity, and overall accuracy were frequently employed, particularly in studies with external validation [64,72]. More recent investigations emphasized clinical evaluation metrics. This reflects a trend toward assessing the potential diagnostic utility of the models rather than relying solely on image similarity measures.

As detailed in Table 2, reported SSIM values were consistently high (>0.97), with Lim et al. [66] and Prieto Canalejo et al. [61] reporting values near 0.99, suggesting strong voxel-level agreement with CT-based references. NMSE values ranged from ~0.5% in Chen et al. [74] to ~3.6% in Shi et al. [65], while PSNR values typically exceeded 30 dB. Torkaman et al. [70] reported a PSNR of ~36.4 dB.

Beyond voxel-level metrics, several studies report effects that map directly to clinical decision making. Shanbhag et al. [64] observed an AUC increase from 0.70 (NAC) to 0.79 with DL-based AC and a rise in normalcy rate from 54.6% to 70.4%, indicating fewer false positives in low-likelihood cohorts while matching CTAC performance (AUC ~0.81). In the multicenter Flurpiridaz-301 analysis, Hagio et al. [72] reported AUC 0.752 vs. 0.717 (DLAC vs. NAC) alongside +6.2% specificity and +4.3% accuracy, reflecting clinically meaningful improvements at constant sensitivity. Prieto Canalejo et al. [61] showed that semi-quantitative SSS categories changed in only ~3% of cases, supporting concordance with reference AC. These results suggest that modest AUC deltas (≈0.03–0.10) can correspond to material gains in specificity/normalcy that reduce artifact-driven positive findings—an outcome with direct clinical impact.

5.4. Do Researchers Utilize Established MPI Quantification Metrics, Such as Summed Stress Score (SSS), Summed Difference Score (SDS), Summed Rest Score (SRS), or Total Perfusion Deficit (TPD) to Evaluate the Quality of DL-Generated AC Images?

The use of established MPI quantification metrics such as SSS, SDS, SRS, and TPD varied across studies. TPD was the most frequently reported clinical metric used in multiple studies, particularly those emphasizing external validation [64,72] and clinical applicability. SSS and SRS were occasionally employed, for example, in the evaluation of FA-ACNet [67], but were not consistently reported. SDS was rarely analyzed independently and was often absent even when other scores were utilized. Several studies, particularly those focused on image-based DL approaches such as voxel-wise cGANs, did not assess any semi-quantitative perfusion scores, relying instead on voxel-level metrics. Thus, while TPD and SSS are sometimes incorporated into evaluations, complete MPI quantification remains underutilized.

This underutilization deserves particular attention, because it raises questions not only about methodological convenience but also about the degree to which emerging computational approaches align with clinical reality.

SSS, SDS, and SRS are underreported in AC studies due to the dependence of these metrics on high-quality, clinically validated ground truth. DL models rely on CT-based attenuation correction as the reference standard. However, CT-based correction itself is not perfect. Misregistration due to respiratory or cardiac motion, beam-hardening artifacts from implanted devices, and anatomical mismatches between emission and transmission data can introduce distortions. Any systematic error is inevitably carried over to the evaluation of the DL output.

Another reason for underutilization is that many studies are conceived and conducted primarily by engineers, computer scientists, or imaging physicists, for whom the most intuitive metrics of success are mathematical measures of similarity between reconstructed and reference images. As a result, SSIM, MAE, and PSNR are not only easy to compute but also familiar benchmarks for optimization. By contrast, semi-quantitative scores like SSS or SDS require additional processing steps, clinical databases, and in some cases expert review. They also require collaboration with cardiologists or nuclear medicine physicians, who may not be directly involved in early technical development. As a result, clinical metrics are sometimes treated as secondary or optional, rather than as primary indicators of diagnostic value.

Nevertheless, the limited use of these metrics represents a missed opportunity. Unlike SSIM or PSNR, which quantify image fidelity in abstract terms, SSS, SDS, and SRS have direct and well-established clinical significance. SSS provides a measure of the overall perfusion defect under stress, SRS describes the extent of fixed defects at rest, and SDS captures the reversible component, which is highly correlated with ischemia and hence with the potential benefit of revascularization. These metrics predict adverse outcomes such as myocardial infarction or cardiac death, and they underpin treatment decision-making in everyday practice [76]. Their incorporation into the evaluation of DL models would therefore fill the gap between technical performance and patient-level clinical impact. An algorithm demonstrating high voxel-level similarity but failing to reproduce the correct SDS value in a patient with inducible ischemia might appear successful but is clinically misleading. Conversely, a model that slightly underperforms on SSIM but manages to preserve SDS and SSS values would have far greater clinical utility.

Another important consideration is the role of these metrics in standardizing evaluation across studies. Because SSS, SRS, and SDS are widely recognized by guidelines and used globally [77], they could serve as a global standard for comparing the clinical validity of different DL approaches.

It is also worth noting that the apparent marginalization of these metrics reflects practical barriers rather than inherent limitations. Some studies have attempted to incorporate TPD or SSS into their validation, showing that it can be done when appropriate infrastructure and clinical collaboration exist. These efforts highlight that the underutilization is not always due to irrelevance but rather to the additional resources required.

5.5. What Are the Typical Sizes and Compositions of Training and Validation Datasets Reported in the Literature?

To provide a more complete overview of the selected studies, the dataset sizes, data splits, metrics, and scanner types are summarized in Tables 1 and 2. The datasets varied widely, ranging from fewer than 100 to more than 11,000 cases, with splits generally between 60 and 80% for training and 15–20% for validation/testing.

Dataset sizes reported across the reviewed literature varied considerably. Only a few studies reported datasets exceeding 1000 patients, such as the framework in [64] which included 4886 cases for training. Most studies used medium-sized datasets ranging from 100 to 400 cases, employing simple splits into training, validation, test sets or cross-validation procedures. Some studies, particularly early explorations of GAN-based methods for attenuation mapping, used very small datasets, sometimes fewer than 100 cases [63,65,69]. Data splits generally followed proportions such as 70% training, 15% validation, and 15% testing. Despite the variability in dataset size, relatively few studies engaged in truly large-scale, multi-institutional data collection efforts.

5.6. How Frequently Do Studies Employ Independent External Populations for Additional Testing?

The use of independent external populations for additional testing was found in only two studies [64,72]. Most investigations relied on internal data splits from a single institution, sometimes labeling held-out internal subsets as “external” validation sets, although they did not originate from distinct institutional sources. True independent external validation was achieved in [64], which evaluated its model on external cohorts from the University of Zurich and University of Calgary comprising 604 patients and in [72].

Table 1. Classification of the reviewed studies based on methodological criteria.

No.	First Author	Year	Ref. No.	Study Design	Camera	Tracer	Data Size	DL Method
1	Shi	2024	[66]	Single-center	Discovery NM/CT 670 (GE Healthcare, Milwaukee, WI, USA)	Tl-201	985 studies	U-Net
2	Hagio	2022	[60]	Single-center	Siemens Symbia T16 (Siemens Healthineers, Malvern, PA, USA)	99mTc-Sestamibi	11,532 studies	U-Net
3	Canalejo	2023	[61]	Single-center	Millenium Hawkeye VG SPECT/CT system (GE Healthcare, Milwaukee, WI, USA)	99mTc-sestamibi	studies	U-Net
4	Yang	2021	[69]	Single-center	Discovery NM/CT 570c scanner (GE Healthcare, Milwaukee, WI, USA)	99mTc-tetrofosmin	100 studies	CNN
5	Huxohl	2022	[62]	Single-center	Symbia Intevo (Siemens Healthineers, Hoffman Estates, USA)	N/R	150 studies	GAN, U-Net
6	Shanbhag	2023	[64]	Multi-center	Discovery 570c or Discovery 530c scanner (GE Healthcare, Milwaukee, WI, USA)	99mTc-sestamibi or 99mTc-tetrofosmin	5490 studies	GAN, U-Net
7	Shi	2020	[65]	Single-center	NM/CT 850 (GE Healthcare, Milwaukee, WI, USA)	99mTc-tetrofosmin	65 studies	GAN, U-Net
8	Yang	2025	[67]	Single-center	NM/CT 670 (GE Healthcare, Milwaukee, WI, USA)	99mTc-sestamibi	202 studies	U-Net
9	Torkaman	2021	[70]	Single-center	Discovery NM 570c (GE Healthcare, Milwaukee, WI, USA)	99mTc-tetrofosmin	100 studies	GAN, U-Net
10	Hagio	2023	[72]	Multi-center	Multiple	99mTc-sestamibi or 99mTc-tetrofosmin	722 studies	CNN
11	Chen	2022	[68]	Single-center	Dedicated SPECT: Alcyone Discovery NM/CT 570c (GE Healthcare, Milwaukee, WI, USA) General SPECT: NM/CT 850c	99mTc-tetrofosmin	Dedicated SPECT: 270 studies General SPECT: 400 studies	U-Net, CNN
12	Torkaman	2022	[71]	Single-center	Discovery NM/CT 570c (GE Healthcare, Milwaukee, WI, USA)	99mTc-tetrofosmin	100 studies	U-Net, GAN
13	Mostafapour	2022	[63]	Single-center	Discovery NM/CT 670 (GE Healthcare, Milwaukee, WI, USA)	99mTc-sestamibi	99 studies	U-Net, CNN
14	Chen	2024	[73]	Single-center	Philips BrightView (Philips, Amsterdam, The Netherlands)	99mTc-sestamibi	1517 studies	U-Net

Table 1. *Cont.*

No.	First Author	Year	Ref. No.	Study Design	Camera	Tracer	Data Size	DL Method
15	Chen	2022	[74]	Single-center	Discovery NM/CT 570c (GE Healthcare, Milwaukee, WI, USA)	99mTc-tetrofosmin	172 studies	CNN
16	Ochoa-Figueroa	2024	[75]	Single-center	D-SPECT (Spectrum Dynamics, Caesarea, Israel)	N/R	300 studies	U-Net

Table 2. Summary of performance metrics. N/R: Not Reported; CV: Cross Validation.

Study	DL Method	Data Split	Metrics	Use of Patient-Specific Features	Has Visual Assessment	Uses Quantitative Analysis (Clinical Metrics)	Uses External Dataset
[66]	U-Net	80%:20%	MAE: 0.003 SSIM: 0.99	✗	✓	✓	✗
[60]	U-Net	60%:20%:20%	R^2: 0.85	✗	✓	✓	✗
[61]	U-Net	320:66	MSSIM: 0.97 ± 0.001 NMAE: 3.08 ± 1.26 (%)	✗	✓	✓	✗
[69]	CNN	10-fold CV	R^2: 0.91 Segmental error: 10%	✗	✓	✗	✗
[62]	GAN, U-Net	70%:15%:15%	NMAE: 0.020 ± 0.007	✗	✓	✗	✗
[64]	GAN, U-Net	4886 train (one site)–604 test (other sites)	Median Absolute Error in TPD: 1.2	✗	✓	✓	✓
[65]	GAN, U-Net	40:25	NMAE: 3.60% ± 0.85%	✗	✓	✗	✗
[67]	U-Net	5-fold CV on 167 studies—25 test	5-fold CV: MSE: $16.94 \pm 2.03 \times 10^{-6}$ SSIM: 0.9955 PSNR: 43.73 ± 0.50 Test set: MSE: 11.98×10^{-6} SSIM: 0.9976 PSNR: 45.54	✗	✓	✓	✗
[70]	GAN, U-Net	5-fold CV	NRMSE: 0.1410 ± 0.0768 PSNR: 36.3823 ± 3.7424 SSIM: 0.9949 ± 0.0043	✗	✓	✗	✗
[72]	CNN	722 test (the model was trained with 11,532 studies of [60])	No reported comparison metrics. AUC was 0.752 for identification of obstructive stenosis using the model's AC images	✗	✓	✓	✓

Table 2. *Cont.*

Study	DL Method	Data Split	Metrics	Use of Patient-Specific Features	Has Visual Assessment	Uses Quantitative Analysis (Clinical Metrics)	Uses External Dataset
[68]	U-Net, CNN	Dedicated SPECT: 100:20:150 General SPECT: 240:60:100	Dedicated SPECT: NMSE: 1.20 ± 0.72% APE: 3.24 ± 2.79% $R^2 = 0.9499$ General SPECT: NMSE: 2.57 ± 1.06%	✓	✓	✗	✗
[71]	U-Net, GAN	leave-one-subject-out cross-validation	NRMSE: 0.135 ± 0.064 PSNR: 36.615 ± 3.45 SSIM: 0.995 ± 0.004	✗	✓	✗	✗
[63]	U-Net, CNN	99:19	ME: −4.41 ± 11.8 SSIM: 0.98 ± 0.05	✗	✓	✓	✗
[73]	U-Net	1131:386	NMSE: 0.5%	✗	✓	✗	✗
[74]	CNN	100:30:42	NMSE: 2.01 ± 1.01%	✓	✓	✗	✗
[75]	U-Net	300 studies (the model is pretrained by the vendor)	No reported comparison metrics.	✗	✓	✗	✗

6. Clinical Applicability and Challenges

The review revealed direct and indirect strategies to transform NAC to AC using DL. Direct image-to-image approaches translate NAC inputs into AC images in image space, requiring no changes to the reconstruction pipeline and yielding fast inference [78]. Their limitations are reduced interpretability and potential global count bias. Indirect approaches predict attenuation (μ) maps that are then used within iterative reconstruction (e.g., OSEM), preserving physics consistency and enabling quality control on the μ-map. However, they can be sensitive to truncated fields-of-view and require re-reconstruction.

Variability in reconstruction pipelines (iterations/subsets, post-filters, resolution recovery, scatter correction) can dominate apparent performance gains. Because many DL-AC models are trained and tested under a single reconstruction recipe, improvements may not persist when parameters are altered within reasonable clinical bounds. We advocate sensitivity analyses across a grid of reconstruction settings and explicit reporting of all parameters. Demonstrating that segmental bias and patient-level indices remain stable across these settings would strengthen claims of robustness and facilitate translation across sites.

Across comparative evaluations, indirect μ-map methods tended to achieve lower NMSE/APE and better segmental agreement (e.g., DuRDN vs. U-Net), while state-of-the-art direct cGANs matched clinical endpoints such as AUC/TPD on external cohorts.

Despite evidence of each effectiveness, the clinical adoption of DL-based attenuation correction methods in SPECT MPI remains limited.

Across studies, 3D attention-augmented U-Nets and cGANs consistently reduced attenuation artifacts and closely matched CT-AC on patient-level indices when evaluated beyond the training site. In contrast, μ-map prediction pipelines tended to excel on voxel and segment level agreement, highlighting their strength for fine-grained spatial fidelity rather than whole-patient metrics. Dataset scale and diversity played a major role, with very large single-center collections delivering strong internal performance but offering limited coverage across vendors and acquisition protocols. Smaller multi-site evaluations provided a better probe of real-world generalizability by exposing models to varied scanners, patient populations, and workflow differences. Persistent gaps remained, including scarce external validation, heterogeneous reporting metrics that hinder direct comparison across papers, and a dependence on CT-derived targets that can propagate CT errors into the downstream deep learning outputs. Addressing these gaps will require standardized benchmarks, broader cross-institutional testing, and reference strategies that do not simply inherit limitations from CT ground truth.

A brief summary of limitations can be found in Table 3. The key limitation is the lack of external validation in most reviewed studies. Internal validation alone cannot ensure consistent performance across diverse patient populations, imaging systems, and institutional protocols. Studies have mitigated limited paired datasets through augmentation and by leveraging multi window inputs that encode scatter information. In practice, these mitigations only partially substitute for genuine diversity in paired data. Augmentation cannot emulate vendor physics, tracer kinetics, or motion patterns. On the other hand, synthetic μ-maps and transfer learning are sensitive to shifts in reconstruction or camera geometry and cross-center splits without protocol heterogeneity can overstate robustness. As such, interpretability artifacts (e.g., over-correction in the inferior wall) may go unnoticed, and reproducibility across institutions remains uncertain without standardized reporting and shared benchmarks.

Table 3. Key gaps in the literature.

In Short	Description
External Evaluation	Most studies rely only on internal validation; very few use independent external cohorts, making generalizability uncertain.
Integration	No models have been prospectively integrated or tested in routine clinical workflows, so real-world usability is unproven.
Compatibility	Models are often trained on data from a single vendor (e.g., GE cameras), and robustness across different scanner types (Siemens, CZT, IQ·SPECT) is unknown.
Multi-center studies	There is a lack of large-scale, multi-institutional collaborations to test robustness across diverse patient populations and acquisition protocols.
CT error propagation	Since CT-based AC is used as ground truth, errors from misregistration or artifacts may be reproduced by DL models. Physics-informed methods could mitigate this but are underexplored.
Non-image data	Patient-specific features (e.g., BMI, gender, chest size) are rarely incorporated, despite potential benefits for contextual accuracy.
Standardized metrics	No consensus exists on evaluation metrics; studies use heterogeneous criteria (SSIM, PSNR, AUC, TPD, etc.), hindering comparisons and clinical adoption.

Multi-center studies that test model robustness across different scanner types, acquisition protocols, and patient demographics, are needed. Without such efforts covering multiple hospitals, it is difficult to ensure consistent diagnostic performance and reliability.

None of the studies reported prospective integration or evaluation of these models in routine clinical settings. As a result, real-world clinical performance, user acceptance, and integration into decision-making workflows remain untested. The absence of clinical trials or longitudinal outcome studies restricts our understanding of how these tools might influence diagnostic and therapeutic decisions [79–81]. Also, existing DL models have not been tested for compatibility with commercial SPECT systems and software. AUC improvements are informative but insufficient to establish clinical utility. Future evaluations should include multi-reader, multi-case (MRMC) ROC or JAFROC studies that quantify inter-reader variability and reader–algorithm interaction. Stratified case sets that emphasize challenging subgroups (e.g., high BMI, breast attenuation, diaphragmatic artifacts, implanted devices) will better reflect day-to-day practice. Reporting reader operating points, variance components, and effect sizes allows comparison to customary gains from CT-AC and clarifies whether DL-AC meaningfully alters diagnostic decisions rather than merely improving pixel-level similarity.

Many high-performing networks were trained exclusively on General Electric solid-state cameras. Hence, their robustness on Siemens IQ·SPECT or CZT pinhole systems is unknown [64]. Models focused on pinhole geometry show excellent NMSE on IQ·SPECT scanners yet may underperform on parallel-hole systems [62]. Domain shift also arises from radiotracer choice (e.g., Tl-201 versus ^{99m}Tc) and from protocol factors such as stress/rest gating, none of which are routinely represented in current training sets.

An overlooked concern is that of error propagation [82] from the CT ground truth. If the reference images are subject to misregistration or beam-hardening artifacts, DL models will likely reproduce them. Physics-informed losses [83] that penalize violations of mass conservation or impose Beer–Lambert constraints [84] could mitigate this, but have yet to be systematically explored.

Most proposed models ignore non-image clinical data, like patient-specific characteristics (BMI, gender, myocardial perimeter), which, if properly incorporated, might enhance performance and contextual understanding.

The creation of an open, vendor-neutral repository containing paired NAC/CT-AC volumes from at least three camera types, two tracers, and 5000+ patients would catalyze development of robust DL models and will serve as global reference. Each case should include clinical outcomes (ICA or FFR) plus demographic variables such as BMI and chest circumference, reflecting the enhanced-feature DuRDN study [74]. Evaluation should be based on four levels. First, voxel-wise performance metrics, such as SSIM. Second, segment-level evaluation using the 17-segment polar map with per-segment bias limits and quantification metrics such as SSS, SRS, SDS. Third, patient-level ROC for obstructive-CAD detection. Last, visual inspection by independent radiologists, cardiologists, or nuclear medicine experts. Because attenuation correction directly alters quantitative readouts, calibration and decision utility should be reported alongside discrimination. Well-calibrated segmental scores and TPD ensure that identical thresholds imply comparable downstream actions before and after DL-AC. Decision-curve analysis and net benefit quantify whether a model improves patient-level decisions across plausible threshold ranges, while reclassification indices test if patients are moved into more appropriate risk strata. Small AUC gains may correspond to large changes in avoidable catheterizations or repeat testing when thresholds are explicit and calibrated.

To achieve this, a coordinated multi-center collaboration would be required, ideally involving academic hospitals, industry partners, and professional societies. Practical steps include establishing standardized data-sharing agreements and federated governance models that ensure patient privacy while enabling research access. Anonymization pipelines and de-identification tools must be integrated into the workflow to remove personally identifiable information while retaining critical clinical and acquisition metadata. Harmonized annotation protocols and uniform formats for NAC/CT-AC pairs would also be essential to guarantee interoperability across vendors. In addition, cloud-based repositories with controlled access could facilitate secure sharing, while federated learning frameworks may allow models to be trained collaboratively without direct data transfer.

Finally, the current methodology lacks standardized performance metrics, evaluation frameworks, or a protocol for model selection and testing. While metrics like SSIM, PSNR, and AUC are commonly reported, their clinical relevance varies, and study comparisons remain inconsistent.

Cardiologists seldom find PSNR or SSIM persuasive, because what they need is an explanation of where and why the algorithm altered the uptake. DeepAC's voxel-level change maps reduced median correction from 9.4 to 2.4 counts compared with CT-AC, providing a visual audit trail [64]. Building on that concept, saliency-guided U-Nets have begun to overlay attention heat-maps on polar plots, highlighting diaphragmatic or breast attenuation that drives the correction. Uncertainty quantification offers a complementary safety net: Monte-Carlo dropout applied to DuRDN outputs revealed 3-to-4× higher predictive variance in right-ventricular myocardium, alerting readers to potential artifacts [68].

A standardized evaluation framework should cover four areas [85]. Technical fidelity at the voxel level can be assessed with SSIM, PSNR, NMSE, NMAE, and correlation coefficients, supported by histogram analysis of voxel intensity distributions in myocardium versus background and Bland–Altman plots for agreement with CT-based attenuation correction. Anatomical or segmental accuracy should be evaluated using 17-segment AHA polar map bias or mean absolute percentage error per segment, correlation coefficients between DL-AC and CT-AC at the segment level, and stratification of errors by anatomical regions such as inferior, anterior, or lateral walls, with statistics presented alongside 95% limits of agreement. Clinical validity requires the use of established myocardial perfusion scores including SSS, SRS, SDS, and TPD. Generalizability and robustness should be demonstrated through external validation on independent datasets from at least one external institution, testing across different scanner vendors such as GE, Siemens, and CZT IQ·SPECT, and evaluation across tracers such as technetium-99m and thallium-201. Subgroup analyses should cover patient characteristics including body mass index, sex, chest circumference, and the presence of implanted devices. Models should further include uncertainty quantification through predictive variance maps and voxel-level change analysis [86]. At minimum, results from one external test cohort, subgroup analyses for high-BMI and female patients, and uncertainty or explainability maps should be provided. Studies should comply with pre-defined AI-reporting criteria [85,87–93].

7. Conclusions

DL has demonstrated potential to transform attenuation correction in SPECT myocardial perfusion imaging by reducing reliance on CT, improving diagnostic accuracy, and simplifying clinical workflow [94]. Robust DL-AC could reduce false positives from soft-tissue artifacts, expand access where CT hardware or dose budgets are constrained, and simplify workflow. Among existing methods, attention-enhanced U-Nets and conditional GANs show the most promising performance, particularly when validated against independent datasets. However, clinical translation remains constrained by limited external validation, heterogeneous evaluation metrics, and minimal integration of non-image patient data.

A notable finding from this review is the imbalance between very large single-center datasets and small cohorts. Large used datasets enable robust training but are often limited to one scanner type and vendor, whereas smaller studies provide methodological innovation but lack external validation. Similarly, evaluation metrics remain inconsistent across studies. Voxel-level similarity indices (SSIM, PSNR, NMSE) were almost universally reported, but established perfusion scores such as SSS, SRS, or SDS were rarely incorporated. To ensure reliability and adoption, future research must emphasize standardized performance criteria that incorporate established MPI quantification scores (SSS, SRS, SDS, TPD), conduct multi-center validation across diverse scanners and tracers, and explore prospective clinical testing [95].

Investigators should define non-inferiority margins to CT-AC and stratify results by vendor, tracer, BMI, and sex. Public benchmarks at scale through multi-center and multi-vendor reference datasets with harmonized preprocessing and challenge style leaderboards that score methods under identical conditions, are needed. Development should advance privacy-preserving generalization using federated learning, domain adaptation, and harmonization to leverage diverse sites without moving data, and external testing must use held-out institutions [96].

Finally, for tools that directly modify quantitative readouts, interpretability (change-maps, saliency on polar plots) and calibrated uncertainty are essential for reader trust and safety checks.

Author Contributions: Conceptualization, I.D.A.; methodology, I.D.A. and N.I.P.; validation, I.D.A., N.I.P., E.I.P. and D.J.A.; investigation, I.D.A.; data curation, I.D.A. and N.I.P.; writing—original draft preparation, I.D.A.; writing—review and editing, N.I.P., D.J.A. and E.I.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	**Full form**
AC	Attenuation Correction
ACM	Attenuation Correction Map
APE	Absolute Percentage Error
AUC	Area Under the Receiver Operating Characteristic Curve
BMI	Body Mass Index
CAD	Coronary Artery Disease
CI	Confidence Interval
CNN	Convolutional Neural Network
CT	Computed Tomography
CTAC	CT-based Attenuation Correction
CTA	Computed Tomography Angiography
CZT	Cadmium–Zinc–Telluride
cGAN	Conditional Generative Adversarial Network
CycleGAN	Cycle-Consistent Generative Adversarial Network
DeepAC	Deep Learning-based Attenuation Correction
DL	Deep Learning
DLAC	Deep Learning Attenuation Correction
DLACS	Deep Learning Attenuation Correction Software
DuRDN	Dual Squeeze-and-Excitation Residual Dense Network
FA-ACNet	Feature-Aligned Attenuation Correction Network
FDA	Food and Drug Administration

FOV	Field of View
GAN	Generative Adversarial Network
GENAC	Generated Attenuation-Corrected
ICA	Invasive Coronary Angiography
ICC	Intraclass Correlation Coefficient
IQ·SPECT	Siemens IQ·SPECT Collimator System
LEHR	Low-Energy High-Resolution
LoA	Limits of Agreement
MAE	Mean Absolute Error
ME	Mean Error
MPI	Myocardial Perfusion Imaging
MSSIM	Mean Structural Similarity Index
MSE	Mean Square Error
NMAE	Normalized Mean Absolute Error
NMSE	Normalized Mean Square Error
NAC	Non-Attenuation Corrected
NRMSE	Normalized Root Mean Square Error
PET	Positron Emission Tomography
PSNR	Peak Signal-to-Noise Ratio
PRAC	Post-Reconstruction Attenuation Correction
ResNet	Residual Neural Network
ResUNet	Residual U-Net
ROC	Receiver Operating Characteristic
R^2	Coefficient of Determination
SDS	Summed Difference Score
SRS	Summed Rest Score
SPECT	Single-Photon Emission Computed Tomography
SSS	Summed Stress Score
SSIM	Structural Similarity Index
TPD	Total Perfusion Deficit

References

1. Huck, D.M.; Weber, B.; Schreiber, B.; Pandav, J.; Parks, S.; Hainer, J.; Brown, J.M.; Divakaran, S.; Blankstein, R.; Dorbala, S.; et al. Comparative Effectiveness of PET and SPECT MPI for Predicting Cardiovascular Events After Kidney Transplant. *Circ. Cardiovasc. Imaging* **2024**, *17*, e015858. [CrossRef] [PubMed]
2. Brindis, R.G.; Douglas, P.S.; Hendel, R.C.; Peterson, E.D.; Wolk, M.J.; Allen, J.M.; Patel, M.R.; Raskin, I.E.; Hendel, R.C.; Bateman, T.M.; et al. ACCF/ASNC Appropriateness Criteria for Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging (SPECT MPI). *J. Am. Coll. Cardiol.* **2005**, *46*, 1587–1605. [CrossRef] [PubMed]
3. Singh, B.; Bateman, T.M.; Case, J.A.; Heller, G. Attenuation Artifact, Attenuation Correction, and the Future of Myocardial Perfusion SPECT. *J. Nucl. Cardiol.* **2007**, *14*, 153–164. [CrossRef] [PubMed]
4. Pazhenkottil, A.P.; Ghadri, J.-R.; Nkoulou, R.N.; Wolfrum, M.; Buechel, R.R.; Küest, S.M.; Husmann, L.; Herzog, B.A.; Gaemperli, O.; Kaufmann, P.A. Improved Outcome Prediction by SPECT Myocardial Perfusion Imaging After CT Attenuation Correction. *J. Nucl. Med.* **2011**, *52*, 196–200. [CrossRef]
5. Sharma, P.; Patel, C.D.; Karunanithi, S.; Maharjan, S.; Malhotra, A. Comparative Accuracy of CT Attenuation-Corrected and Non–Attenuation-Corrected SPECT Myocardial Perfusion Imaging. *Clin. Nucl. Med.* **2012**, *37*, 332–338. [CrossRef]
6. Hsieh, J.; Flohr, T. Computed Tomography Recent History and Future Perspectives. *J. Med. Imaging* **2021**, *8*, 052109. [CrossRef]
7. Shao, W.; Rowe, S.P.; Du, Y. Artificial Intelligence in Single Photon Emission Computed Tomography (SPECT) Imaging: A Narrative Review. *Ann. Transl. Med.* **2021**, *9*, 820. [CrossRef]
8. Berker, Y.; Li, Y. Attenuation Correction in Emission Tomography Using the Emission Data—A Review. *Med. Phys.* **2016**, *43*, 807–832. [CrossRef]
9. Hutton, B.F. The Origins of SPECT and SPECT/CT. *Eur. J. Nucl. Med. Mol. Imaging* **2014**, *41*, 3–16. [CrossRef]
10. Patton, J.A.; Turkington, T.G. SPECT/CT Physical Principles and Attenuation Correction. *J. Nucl. Med. Technol.* **2008**, *36*, 1–10. [CrossRef]

11. Welch, A.; Clack, R.; Natterer, F.; Gullberg, G.T. Toward Accurate Attenuation Correction in SPECT without Transmission Measurements. *IEEE Trans. Med. Imaging* **1997**, *16*, 532–541. [CrossRef] [PubMed]
12. Tadesse, G.F.; Geramifar, P.; Tegaw, E.M.; Ay, M.R. Techniques for Generating Attenuation Map Using Cardiac SPECT Emission Data Only: A Systematic Review. *Ann. Nucl. Med.* **2019**, *33*, 1–13. [CrossRef]
13. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
14. Hajianfar, G.; Gharibi, O.; Sabouri, M.; Mohebi, M.; Amini, M.; Yasemi, M.J.; Chehreghani, M.; Maghsudi, M.; Mansouri, Z.; Edalat-Javid, M.; et al. Artificial Intelligence-Powered Coronary Artery Disease Diagnosis from SPECT Myocardial Perfusion Imaging: A Comprehensive Deep Learning Study. *Eur. J. Nucl. Med. Mol. Imaging* **2025**, *52*, 3019–3035. [CrossRef] [PubMed]
15. Balaji, V.; Song, T.-A.; Malekzadeh, M.; Heidari, P.; Dutta, J. Artificial Intelligence for PET and SPECT Image Enhancement. *J. Nucl. Med.* **2024**, *65*, 4–12. [CrossRef]
16. Kawakubo, M.; Nagao, M.; Kaimoto, Y.; Nakao, R.; Yamamoto, A.; Kawasaki, H.; Iwaguchi, T.; Matsuo, Y.; Kaneko, K.; Sakai, A.; et al. Deep Learning Approach Using SPECT-to-PET Translation for Attenuation Correction in CT-Less Myocardial Perfusion SPECT Imaging. *Ann. Nucl. Med.* **2024**, *38*, 199–209. [CrossRef]
17. Kwon, K.; Oh, D.; Kim, J.H.; Yoo, J.; Lee, W.W. Deep-Learning-Based Attenuation Map Generation in Kidney Single Photon Emission Computed Tomography. *EJNMMI Phys.* **2024**, *11*, 84. [CrossRef]
18. Jia, Y.; Li, Z.; Akhavanallaf, A.; Fessler, J.A.; Dewaraja, Y.K. 90Y SPECT Scatter Estimation and Voxel Dosimetry in Radioembolization Using a Unified Deep Learning Framework. *EJNMMI Phys.* **2023**, *10*, 82. [CrossRef]
19. Le, T.D.; Shitiri, N.C.; Jung, S.-H.; Kwon, S.-Y.; Lee, C. Image Synthesis in Nuclear Medicine Imaging with Deep Learning: A Review. *Sensors* **2024**, *24*, 8068. [CrossRef]
20. Sengupta, P.P.; Tokodi, M. Automated Interpretation of Myocardial Perfusion Images. *JACC Cardiovasc. Imaging* **2022**, *15*, 1103–1106. [CrossRef]
21. Saboury, B.; Bradshaw, T.; Boellaard, R.; Buvat, I.; Dutta, J.; Hatt, M.; Jha, A.K.; Li, Q.; Liu, C.; McMeekin, H.; et al. Artificial Intelligence in Nuclear Medicine: Opportunities, Challenges, and Responsibilities Toward a Trustworthy Ecosystem. *J. Nucl. Med.* **2023**, *64*, 188–196. [CrossRef] [PubMed]
22. Kusumoto, D.; Akiyama, T.; Hashimoto, M.; Iwabuchi, Y.; Katsuki, T.; Kimura, M.; Akiba, Y.; Sawada, H.; Inohara, T.; Yuasa, S.; et al. A Deep Learning-Based Automated Diagnosis System for SPECT Myocardial Perfusion Imaging. *Sci. Rep.* **2024**, *14*, 13583. [CrossRef] [PubMed]
23. Tadesse, G.F.; Geramifar, P.; Abbasi, M.; Tsegaw, E.M.; Amin, M.; Salimi, A.; Mohammadi, M.; Teimourianfard, B.; Ay, M.R. Attenuation Correction for Dedicated Cardiac SPECT Imaging Without Using Transmission Data. *Mol. Imaging Radionucl. Ther.* **2023**, *32*, 42–53. [CrossRef]
24. LeCun, Y.; Kavukcuoglu, K.; Farabet, C. Convolutional networks and applications in vision. In Proceedings of the 2010 IEEE International Symposium on Circuits and Systems, Paris, France, 30 May–2 June 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 253–256.
25. Dong, S.; Wang, P.; Abbas, K. A Survey on Deep Learning and Its Applications. *Comput. Sci. Rev.* **2021**, *40*, 100379. [CrossRef]
26. Elasri, M.; Elharrouss, O.; Al-Maadeed, S.; Tairi, H. Image Generation: A Review. *Neural Process. Lett.* **2022**, *54*, 4609–4646. [CrossRef]
27. Baraheem, S.S.; Le, T.-N.; Nguyen, T.V. Image Synthesis: A Review of Methods, Datasets, Evaluation Metrics, and Future Outlook. *Artif. Intell. Rev.* **2023**, *56*, 10813–10865. [CrossRef]
28. Apostolopoulos, D.J.; Savvopoulos, C. What Is the Benefit of CT-Based Attenuation Correction in Myocardial Perfusion SPET? *Hell. J. Nucl. Med.* **2016**, *19*, 89–92.
29. Verberne, H.J.; Acampa, W.; Anagnostopoulos, C.; Ballinger, J.; Bengel, F.; De Bondt, P.; Buechel, R.R.; Cuocolo, A.; Van Eck-Smit, B.L.F.; Flotats, A.; et al. EANM Procedural Guidelines for Radionuclide Myocardial Perfusion Imaging with SPECT and SPECT/CT: 2015 Revision. *Eur. J. Nucl. Med. Mol. Imaging* **2015**, *42*, 1929–1940. [CrossRef]
30. Bouchareb, Y.; AlSaadi, A.; Zabah, J.; Jain, A.; Al-Jabri, A.; Phiri, P.; Shi, J.Q.; Delanerolle, G.; Sirasanagandla, S.R. Technological Advances in SPECT and SPECT/CT Imaging. *Diagnostics* **2024**, *14*, 1431. [CrossRef]
31. Danad, I.; Raijmakers, P.G.; Driessen, R.S.; Leipsic, J.; Raju, R.; Naoum, C.; Knuuti, J.; Mäki, M.; Underwood, R.S.; Min, J.K.; et al. Comparison of Coronary CT Angiography, SPECT, PET, and Hybrid Imaging for Diagnosis of Ischemic Heart Disease Determined by Fractional Flow Reserve. *JAMA Cardiol.* **2017**, *2*, 1100. [CrossRef]
32. Abbott, B.G.; Case, J.A.; Dorbala, S.; Einstein, A.J.; Galt, J.R.; Pagnanelli, R.; Bullock-Palmer, R.P.; Soman, P.; Wells, R.G. Contemporary Cardiac SPECT Imaging—Innovations and Best Practices: An Information Statement from the American Society of Nuclear Cardiology. *Circ. Cardiovasc. Imaging* **2018**, *11*, e000020. [CrossRef]
33. Chen, X.; Liu, C. Deep-Learning-Based Methods of Attenuation Correction for SPECT and PET. *J. Nucl. Cardiol.* **2023**, *30*, 1859–1878. [CrossRef]

34. Israel, O.; Pellet, O.; Biassoni, L.; De Palma, D.; Estrada-Lobato, E.; Gnanasegaran, G.; Kuwert, T.; La Fougère, C.; Mariani, G.; Massalha, S.; et al. Two Decades of SPECT/CT—The Coming of Age of a Technology: An Updated Review of Literature Evidence. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 1990–2012. [CrossRef]
35. Seo, Y.; Mari, C.; Hasegawa, B.H. Technological Development and Advances in Single-Photon Emission Computed Tomography/Computed Tomography. *Semin. Nucl. Med.* **2008**, *38*, 177–198. [CrossRef]
36. Cheng, Z.; Chen, P.; Yan, J. A Review of State-of-the-Art Resolution Improvement Techniques in SPECT Imaging. *EJNMMI Phys.* **2025**, *12*, 9. [CrossRef]
37. Van Audenhaege, K.; Van Holen, R.; Vandenberghe, S.; Vanhove, C.; Metzler, S.D.; Moore, S.C. Review of SPECT Collimator Selection, Optimization, and Fabrication for Clinical and Preclinical Imaging. *Med. Phys.* **2015**, *42*, 4796–4813. [CrossRef]
38. Shiba, S.; Sagara, H. Passive Gamma Emission Tomography with Ordered Subset Expectation Maximization Method. *Ann. Nucl. Energy* **2021**, *150*, 107823. [CrossRef]
39. Trevisan, A.C.; Raed, M.D.; Tumas, V.; Alexandre-Santos, L.; Pitella, F.A.; Itikawa, E.N.; Silvah, J.H.; Kato, M.; Martinez, E.Z.; Achcar, J.A.; et al. Comparison between OSEM and FBP Reconstruction Algorithms for the Qualitative and Quantitative Interpretation of Brain DAT-SPECT Using an Anthropomorphic Striatal Phantom: Implications for the Practice. *Res. Biomed. Eng.* **2020**, *36*, 77–88. [CrossRef]
40. Dong, X.; Lei, Y.; Wang, T.; Higgins, K.; Liu, T.; Curran, W.J.; Mao, H.; Nye, J.A.; Yang, X. Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging. *Phys. Med. Biol.* **2020**, *65*, 055011. [CrossRef] [PubMed]
41. Yang, J.; Sohn, J.H.; Behr, S.C.; Gullberg, G.T.; Seo, Y. CT-Less Direct Correction of Attenuation and Scatter in the Image Space Using Deep Learning for Whole-Body FDG PET: Potential Benefits and Pitfalls. *Radiol. Artif. Intell.* **2021**, *3*, e200137. [CrossRef]
42. Arabi, H.; Zaidi, H. Deep Learning-Guided Estimation of Attenuation Correction Factors from Time-of-Flight PET Emission Data. *Med. Image Anal.* **2020**, *64*, 101718. [CrossRef]
43. Krokos, G.; MacKewn, J.; Dunn, J.; Marsden, P. A Review of PET Attenuation Correction Methods for PET-MR. *EJNMMI Phys.* **2023**, *10*, 52. [CrossRef]
44. Jahangir, R.; Kamali-Asl, A.; Arabi, H.; Zaidi, H. Strategies for Deep Learning-based Attenuation and Scatter Correction of Brain^{18} F-FDG PET Images in the Image Domain. *Med. Phys.* **2024**, *51*, 870–880. [CrossRef]
45. Murata, T.; Yokota, H.; Yamato, R.; Horikoshi, T.; Tsuneda, M.; Kurosawa, R.; Hashimoto, T.; Ota, J.; Sawada, K.; Iimori, T.; et al. Development of Attenuation Correction Methods Using Deep Learning in Brain-perfusion Single-photon Emission Computed Tomography. *Med. Phys.* **2021**, *48*, 4177–4190. [CrossRef] [PubMed]
46. Sakaguchi, K.; Kaida, H.; Yoshida, S.; Ishii, K. Attenuation Correction Using Deep Learning for Brain Perfusion SPECT Images. *Ann. Nucl. Med.* **2021**, *35*, 589–599. [CrossRef] [PubMed]
47. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
48. LeCun, Y.; Bengio, Y. Convolutional Networks for Images, Speech, and Time Series. In *The Handbook of Brain Theory and Neural Networks*; MIT Press: Cambridge, MA, USA, 1995; Volume 3361.
49. Apostolopoulos, I.D.; Papandrianos, N.I.; Apostolopoulos, D.J.; Papageorgiou, E. Between Two Worlds: Investigating the Intersection of Human Expertise and Machine Learning in the Case of Coronary Artery Disease Diagnosis. *Bioengineering* **2024**, *11*, 957. [CrossRef] [PubMed]
50. Apostolopoulos, D.J.; Apostolopoulos, I.D.; Papathanasiou, N.D.; Spyridonidis, T.; Panayiotakis, G.S. Explainable Artificial Intelligence Method (ParaNet+) Localises Abnormal Parathyroid Glands in Scintigraphic Scans of Patients with Primary Hyperparathyroidism. *Algorithms* **2023**, *16*, 435. [CrossRef]
51. Feleki, A.; Apostolopoulos, I.D.; Moustakidis, S.; Papageorgiou, E.I.; Papathanasiou, N.; Apostolopoulos, D.; Papandrianos, N. Explainable Deep Fuzzy Cognitive Map Diagnosis of Coronary Artery Disease: Integrating Myocardial Perfusion Imaging, Clinical Data, and Natural Language Insights. *Appl. Sci.* **2023**, *13*, 11953. [CrossRef]
52. Feleki, A.; Apostolopoulos, I.D.; Papageorgiou, K.; Papageorgiou, E.I.; Apostolopoulos, D.J.; Papandrianos, N.I. A Fuzzy Cognitive Map Learning Approach for Coronary Artery Disease Diagnosis in Nuclear Medicine. In *Fuzzy Logic and Technology, and Aggregation Operators*; Massanet, S., Montes, S., Ruiz-Aguilera, D., González-Hidalgo, M., Eds.; Lecture Notes in Computer Science; Springer Nature: Cham, Switzerland, 2023; Volume 14069, pp. 14–25. ISBN 978-3-031-39964-0.
53. Papandrianos, N.I.; Apostolopoulos, I.D.; Feleki, A.; Apostolopoulos, D.J.; Papageorgiou, E.I. Deep Learning Exploration for SPECT MPI Polar Map Images Classification in Coronary Artery Disease. *Ann. Nucl. Med.* **2022**, *36*, 823–833. [CrossRef]
54. Samaras, A.-D.; Moustakidis, S.; Apostolopoulos, I.D.; Papageorgiou, E.; Papandrianos, N. Uncovering the Black Box of Coronary Artery Disease Diagnosis: The Significance of Explainability in Predictive Models. *Appl. Sci.* **2023**, *13*, 8120. [CrossRef]
55. Domingues, I.; Pereira, G.; Martins, P.; Duarte, H.; Santos, J.; Abreu, P.H. Using Deep Learning Techniques in Medical Imaging: A Systematic Review of Applications on CT and PET. *Artif. Intell. Rev.* **2020**, *53*, 4093–4160. [CrossRef]
56. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385. [CrossRef]

57. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 9351, pp. 234–241. ISBN 978-3-319-24573-7.
58. Ali, S.; Abuhmed, T.; El-Sappagh, S.; Muhammad, K.; Alonso-Moral, J.M.; Confalonieri, R.; Guidotti, R.; Del Ser, J.; Díaz-Rodríguez, N.; Herrera, F. Explainable Artificial Intelligence (XAI): What We Know and What Is Left to Attain Trustworthy Artificial Intelligence. *Inf. Fusion* **2023**, *99*, 101805. [CrossRef]
59. Zhang, Q.; Zhu, S.-C. Visual interpretability for deep learning: A survey. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, 27–39. [CrossRef]
60. Hagio, T.; Poitrasson-Rivière, A.; Moody, J.B.; Renaud, J.M.; Arida-Moody, L.; Shah, R.V.; Ficaro, E.P.; Murthy, V.L. "Virtual" Attenuation Correction: Improving Stress Myocardial Perfusion SPECT Imaging Using Deep Learning. *Eur. J. Nucl. Med. Mol. Imaging* **2022**, *49*, 3140–3149. [CrossRef] [PubMed]
61. Prieto Canalejo, M.; Palau San Pedro, A.; Geronazzo, R.; Minsky, D.; Juárez-Orozco, L.; Namías, M. Synthetic Attenuation Correction Maps for SPECT Imaging Using Deep Learning: A Study on Myocardial Perfusion Imaging. *Diagnostics* **2023**, *13*, 2214. [CrossRef]
62. Huxohl, T.; Patel, G.; Zabel, R.; Burchert, W. Deep Learning Approximation of Attenuation Maps for Myocardial Perfusion SPECT with an IQ SPECT Collimator. *EJNMMI Phys.* **2023**, *10*, 49. [CrossRef]
63. Mostafapour, S.; Gholamiankhah, F.; Maroufpour, S.; Momennezhad, M.; Asadinezhad, M.; Zakavi, S.R.; Arabi, H.; Zaidi, H. Deep Learning-Guided Attenuation Correction in the Image Domain for Myocardial Perfusion SPECT Imaging. *J. Comput. Des. Eng.* **2022**, *9*, 434–447. [CrossRef]
64. Shanbhag, A.D.; Miller, R.J.H.; Pieszko, K.; Lemley, M.; Kavanagh, P.; Feher, A.; Miller, E.J.; Sinusas, A.J.; Kaufmann, P.A.; Han, D.; et al. Deep Learning–Based Attenuation Correction Improves Diagnostic Accuracy of Cardiac SPECT. *J. Nucl. Med.* **2023**, *64*, 472–478. [CrossRef]
65. Shi, L.; Onofrey, J.A.; Liu, H.; Liu, Y.-H.; Liu, C. Deep Learning-Based Attenuation Map Generation for Myocardial Perfusion SPECT. *Eur. J. Nucl. Med. Mol. Imaging* **2020**, *47*, 2383–2395. [CrossRef]
66. Lim, S.; Park, Y.-J.; Lee, S.J.; An, Y.-S.; Yoon, J.-K. Clinical Feasibility of Deep Learning–Based Attenuation Correction Models for Tl-201 Myocardial Perfusion SPECT. *Clin. Nucl. Med.* **2024**, *49*, 397–403. [CrossRef] [PubMed]
67. Yang, P.; Zhang, Z.; Wei, J.; Jiang, L.; Yu, L.; Cai, H.; Li, L.; Guo, Q.; Zhao, Z. Deep Learning-Based CT-Free Attenuation Correction for Cardiac SPECT: A New Approach. *BMC Med. Imaging* **2025**, *25*, 38. [CrossRef]
68. Chen, X.; Zhou, B.; Xie, H.; Shi, L.; Liu, H.; Holler, W.; Lin, M.; Liu, Y.-H.; Miller, E.J.; Sinusas, A.J.; 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* **2022**, *49*, 3046–3060. [CrossRef] [PubMed]
69. Yang, J.; Shi, L.; Wang, R.; Miller, E.J.; Sinusas, A.J.; Liu, C.; Gullberg, G.T.; Seo, Y. Direct Attenuation Correction Using Deep Learning for Cardiac SPECT: A Feasibility Study. *J. Nucl. Med.* **2021**, *62*, 1645–1652. [CrossRef]
70. Torkaman, M.; Yang, J.; Shi, L.; Wang, R.; Miller, E.J.; Sinusas, A.J.; Liu, C.; Gullberg, G.T.; Seo, Y. Direct Image-Based Attenuation Correction Using Conditional Generative Adversarial Network for SPECT Myocardial Perfusion Imaging. In Proceedings of the Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging, Online Only, 12–15 February 2021; Gimi, B.S., Krol, A., Eds.; SPIE: Washington, DC, USA, 2021; p. 27.
71. Torkaman, M.; Yang, J.; Shi, L.; Wang, R.; Miller, E.J.; Sinusas, A.J.; Liu, C.; Gullberg, G.T.; Seo, Y. Data Management and Network Architecture Effect on Performance Variability in Direct Attenuation Correction via Deep Learning for Cardiac SPECT: A Feasibility Study. *IEEE Trans. Radiat. Plasma Med. Sci.* **2022**, *6*, 755–765. [CrossRef]
72. Hagio, T.; Moody, J.B.; Poitrasson-Rivière, A.; Renaud, J.M.; Pierce, L.; Buckley, C.; Ficaro, E.P.; Murthy, V.L. Multi-Center, Multi-Vendor Validation of Deep Learning-Based Attenuation Correction in SPECT MPI: Data from the International Flurpiridaz-301 Trial. *Eur. J. Nucl. Med. Mol. Imaging* **2023**, *50*, 1028–1033. [CrossRef]
73. Chen, Y.; Pretorius, P.H.; Yang, Y.; King, M.A.; Lindsay, C. Investigation of Scatter Energy Window Width and Count Levels for Deep Learning-Based Attenuation Map Estimation in Cardiac SPECT/CT Imaging. *Phys. Med. Biol.* **2024**, *69*, 225009. [CrossRef]
74. Chen, X.; Zhou, B.; Shi, L.; Liu, H.; Pang, Y.; Wang, R.; Miller, E.J.; Sinusas, A.J.; Liu, C. CT-Free Attenuation Correction for Dedicated Cardiac SPECT Using a 3D Dual Squeeze-and-Excitation Residual Dense Network. *J. Nucl. Cardiol.* **2022**, *29*, 2235–2250. [CrossRef]
75. Ochoa-Figueroa, M.; Valera-Soria, C.; Pagonis, C.; Ressner, M.; Norberg, P.; Sanchez-Rodriguez, V.; Frias-Rose, J.; Good, E.; Davidsson, A. Diagnostic Performance of a Novel Deep Learning Attenuation Correction Software for MPI Using a Cardio Dedicated CZT Camera. Experience in the Clinical Practice. *Rev. Esp. Med. Nucl. Imagen Mol. Engl. Ed.* **2024**, *43*, 23–30. [CrossRef]
76. Hachamovitch, R.; Hayes, S.W.; Friedman, J.D.; Cohen, I.; Berman, D.S. Comparison of the Short-Term Survival Benefit Associated with Revascularization Compared with Medical Therapy in Patients with No Prior Coronary Artery Disease Undergoing Stress Myocardial Perfusion Single Photon Emission Computed Tomography. *Circulation* **2003**, *107*, 2900–2907. [CrossRef]

77. Hachamovitch, R.; Berman, D.S.; Shaw, L.J.; Kiat, H.; Cohen, I.; Cabico, J.A.; Friedman, J.; Diamond, G.A. Incremental Prognostic Value of Myocardial Perfusion Single Photon Emission Computed Tomography for the Prediction of Cardiac Death: Differential Stratification for Risk of Cardiac Death and Myocardial Infarction. *Circulation* **1998**, *97*, 535–543. [CrossRef] [PubMed]
78. Yang, F.; Lei, B.; Zhou, Z.; Song, T.-A.; Balaji, V.; Dutta, J. AI in SPECT Imaging: Opportunities and Challenges. *Semin. Nucl. Med.* **2025**, *55*, 294–312. [CrossRef] [PubMed]
79. Miller, R.J.H.; Slomka, P.J. Artificial Intelligence in Nuclear Cardiology: An Update and Future Trends. *Semin. Nucl. Med.* **2024**, *54*, 648–657. [CrossRef] [PubMed]
80. Sciagrà, R.; Valente, S.; Dominietto, M. Artificial Intelligence in Nuclear Cardiology. *J. Clin. Med.* **2025**, *14*, 6416. [CrossRef]
81. Hu, X.; Zhang, H.; Caobelli, F.; Huang, Y.; Li, Y.; Zhang, J.; Shi, K.; Yu, F. The Role of Deep Learning in Myocardial Perfusion Imaging for Diagnosis and Prognosis: A Systematic Review. *iScience* **2024**, *27*, 111374. [CrossRef]
82. Gould, K.L.; Pan, T.; Loghin, C.; Johnson, N.P.; Guha, A.; Sdringola, S. Frequent Diagnostic Errors in Cardiac PET/CT Due to Misregistration of CT Attenuation and Emission PET Images: A Definitive Analysis of Causes, Consequences, and Corrections. *J. Nucl. Med.* **2007**, *48*, 1112–1121. [CrossRef]
83. Gong, H.; Tao, S.; Rajendran, K.; Zhou, W.; McCollough, C.H.; Leng, S. Deep-learning-based Direct Inversion for Material Decomposition. *Med. Phys.* **2020**, *47*, 6294–6309. [CrossRef]
84. Ghosh, S.; Varshney, N.; Al Hasan, M.M.; Roy, A.; Craig, P.; Koppal, S.J.; Dalir, H.; Asadizanjani, N. Exploring Physics-Informed Machine Learning for System Matrix Formulation in x-Ray Imaging Forward Models. In *Proceedings of the Developments in X-Ray Tomography XV*; Müller, B., Wang, G., Eds.; SPIE: San Diego, CA, USA, 2024; p. 66.
85. Collins, G.S.; Moons, K.G.M.; Dhiman, P.; Riley, R.D.; Beam, A.L.; Van Calster, B.; Ghassemi, M.; Liu, X.; Reitsma, J.B.; Van Smeden, M.; et al. TRIPOD+AI Statement: Updated Guidance for Reporting Clinical Prediction Models That Use Regression or Machine Learning Methods. *BMJ* **2024**, *385*, e078378. [CrossRef]
86. Maier-Hein, L.; Reinke, A.; Godau, P.; Tizabi, M.D.; Buettner, F.; Christodoulou, E.; Glocker, B.; Isensee, F.; Kleesiek, J.; Kozubek, M.; et al. Metrics Reloaded: Recommendations for Image Analysis Validation. *Nat. Methods* **2024**, *21*, 195–212. [CrossRef]
87. Moons, K.G.M.; Damen, J.A.A.; Kaul, T.; Hooft, L.; Andaur Navarro, C.; Dhiman, P.; Beam, A.L.; Van Calster, B.; Celi, L.A.; Denaxas, S.; et al. PROBAST+AI: An Updated Quality, Risk of Bias, and Applicability Assessment Tool for Prediction Models Using Regression or Artificial Intelligence Methods. *BMJ* **2025**, *388*, e082505. [CrossRef]
88. Liu, X.; Rivera, S.C.; Moher, D.; Calvert, M.J.; Denniston, A.K. Reporting Guidelines for Clinical Trial Reports for Interventions Involving Artificial Intelligence: The CONSORT-AI Extension. *BMJ* **2020**, *370*, m3164. [CrossRef]
89. Rivera, S.C.; Liu, X.; Chan, A.-W.; Denniston, A.K.; Calvert, M.J. Guidelines for Clinical Trial Protocols for Interventions Involving Artificial Intelligence: The SPIRIT-AI Extension. *BMJ* **2020**, *370*, m3210. [CrossRef]
90. Vasey, B.; Nagendran, M.; Campbell, B.; Clifton, D.A.; Collins, G.S.; Denaxas, S.; Denniston, A.K.; Faes, L.; Geerts, B.; Ibrahim, M.; et al. Reporting Guideline for the Early-Stage Clinical Evaluation of Decision Support Systems Driven by Artificial Intelligence: DECIDE-AI. *Nat. Med.* **2022**, *28*, 924–933. [CrossRef]
91. Mongan, J.; Moy, L.; Kahn, C.E. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiol. Artif. Intell.* **2020**, *2*, e200029. [CrossRef] [PubMed]
92. Norgeot, B.; Quer, G.; Beaulieu-Jones, B.K.; Torkamani, A.; Dias, R.; Gianfrancesco, M.; Arnaout, R.; Kohane, I.S.; Saria, S.; Topol, E.; et al. Minimum Information about Clinical Artificial Intelligence Modeling: The MI-CLAIM Checklist. *Nat. Med.* **2020**, *26*, 1320–1324. [CrossRef] [PubMed]
93. Hernandez-Boussard, T.; Bozkurt, S.; Ioannidis, J.P.A.; Shah, N.H. MINIMAR (MINimum Information for Medical AI Reporting): Developing Reporting Standards for Artificial Intelligence in Health Care. *J. Am. Med. Inform. Assoc.* **2020**, *27*, 2011–2015. [CrossRef] [PubMed]
94. Saboury, B.; Ghesani, M. Artificial Intelligence in Nuclear Medicine: Point—More Reality Than Hype Today. *Am. J. Roentgenol.* **2024**, *223*, e2431288. [CrossRef]
95. Shanbhag, A.D.; Miller, R.J.H.; Lemley, M.; Kavanagh, P.; Liang, J.X.; Marcinkiewicz, A.M.; Builoff, V.; Van Kriekinge, S.; Ruddy, T.D.; Fish, M.B.; et al. General Purpose Deep Learning Attenuation Correction Improves Diagnostic Accuracy of SPECT MPI. *JACC Cardiovasc. Imaging* **2025**, S1936878X25003316. [CrossRef]
96. Gichoya, J.W.; Banerjee, I.; Bhimireddy, A.R.; Burns, J.L.; Celi, L.A.; Chen, L.-C.; Correa, R.; Dullerud, N.; Ghassemi, M.; Huang, S.-C.; et al. AI Recognition of Patient Race in Medical Imaging: A Modelling Study. *Lancet Digit. Health* **2022**, *4*, e406–e414. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.