Advances in Science and Technology
Research Journal

Advances in Science and Technology Research Journal, 2025, 19(11), 76–90
https://doi.org/10.12913/22998624/208545
ISSN 2299-8624, License CC-BY 4.0

Received: 2025.06.24
Accepted: 2025.09.15
Published: 2025.10.01

Computed tomography-positron emission and tomography image reconstruction under compressed sensing constraints

Krzysztof Malczewski[1]

[1] Institute of Information Technology, Warsaw University of Life Sciences, ul. Nowoursynowska 159, Building 34, 02-776 Warsaw, Poland
E-mail: krzysztof_malczewski@sggw.edu.pl

ABSTRACT

A novel generative framework, RED-WGAN, is presented for the reconstruction of high-resolution CT-PET images under conditions of sparse sampling, motion corruption, and low-dose acquisition. The method integrates deformable motion compensation, residual deblurring, and Transformer-based denoising within a unified architecture. Joint priors from structural (CT) and functional (PET) data are combined with Ridgelet-domain sparsity and compressed sensing principles to enable the recovery of anatomically precise and perceptually coherent images from highly degraded inputs. The framework was evaluated using both digital phantom models (Shepp-Logan and Zubal) and clinical in vivo CT-PET datasets encompassing a range of anatomical sites and motion scenarios. RED-WGAN was benchmarked against twelve state-of-the-art super-resolution algorithms under varying levels of data sparsity (20–100%) and motion distortion. Superior performance was consistently observed, with PSNR values reaching 39.03 dB, SSIM up to 0.921, and LPIPS reduced to 0.118. Robust image quality was maintained at compression ratios as low as 40%. Furthermore, the framework achieved the lowest total registration error (1.44 voxels) across all evaluated motion compensation methods. These findings demonstrate that RED-WGAN offers a technically robust and clinically scalable solution for CT-PET image enhancement. Its ability to preserve structural integrity and improve cross-modality alignment under constrained acquisition conditions positions it as a promising tool for applications in low-dose imaging, pediatric diagnostics, rapid scanning protocols, and motion-prone clinical environments.

Keywords: GAN, WGAN, generative, Wasserstein, computed tomography, positron emission, super-resolution image reconstruction, compressive sensing.

INTRODUCTION

Hybrid imaging techniques have become a cornerstone in modern diagnostic radiology, with Computed Tomography–Positron Emission Tomography (CT-PET) emerging as a gold standard in the evaluation of various oncological, neurological, and cardiovascular conditions. CT provides detailed anatomical structure, while PET highlights metabolic and molecular activity through the uptake of radiotracers such as F-FDG. Their integration enables the simultaneous acquisition of functional and structural information, providing a unified and comprehensive perspective for lesion localization, staging, and therapy assessment. Despite its clinical advantages, CT-PET imaging remains constrained by several technical and practical limitations. These include prolonged acquisition times, increased patient radiation exposure, and the presence of motion artifacts caused by respiratory and physiological movement during scan sessions. Furthermore, PET imaging is typically acquired at a much lower resolution than CT, which can lead to spatial mismatch during image fusion, particularly in mobile or deformable organs. Image degradation due to noise, blur, and under-sampling further diminishes diagnostic interpretability, especially in pediatric or low-dose settings where acquisition time and injected activity must be minimized.

In recent years, artificial intelligence and deep learning have been increasingly employed to address these limitations through advanced image reconstruction strategies. Super-resolution (SR) methods based on deep convolutional neural networks (CNNs), residual attention mechanisms, and generative adversarial networks (GANs) [1–4] have shown considerable promise in recovering fine-grained anatomical detail from coarse or sparsely sampled inputs. However, many existing models have not been explicitly designed for hybrid imaging contexts (i.e. here CT-PET image reconstruction), and often rely on ideal acquisition assumptions without addressing multi-modal misalignment, motion distortion, or domain-specific anatomical constraints. To address these challenges, a novel framework termed RED-WGAN (Registration-Enhanced Deformable Wasserstein GAN) is proposed in this work. The method is specifically tailored for CT-PET reconstruction, and integrates a motion-aware deformable registration network within a deep super-resolution architecture. The approach combines compressed sensing acquisition with perceptually guided adversarial learning to reconstruct high-resolution CT-PET images from low-dose, motion-affected, or sparsely sampled data. By leveraging joint priors derived from both CT and PET modalities, the RED-WGAN framework [5] achieves improved structural delineation and accurate metabolic representation.

In addition, a Ridgelet-based sparsity model is incorporated to better exploit the compressibility of CT-PET signals in the transform domain. Compressed sensing principles are embedded into both the acquisition simulation and the network training strategy, thereby enabling reliable reconstructions from as little as 40% of the raw acquisition data. A deformable motion estimation module further corrects for non-rigid anatomical displacement without requiring handcrafted motion vectors, improving cross-modality registration consistency and image quality in dynamic or free-breathing acquisitions.

The proposed model is trained using a composite loss function that combines pixel-wise content loss, perceptual similarity metrics (e.g., LPIPS), and Wasserstein adversarial loss, in order to balance structural accuracy with perceptual realism. The framework is validated on extensive in vivo and phantom datasets using established quantitative metrics such as PSNR, SSIM, RMSE, and total registration error (TRE). The results indicate that RED-WGAN consistently outperforms conventional interpolation, classical deep SR methods, and existing GAN-based models in terms of resolution, artifact reduction, and diagnostic fidelity. The major contributions of this study are summarized as follows:

- A novel GAN-based framework (RED-WGAN) is proposed for CT-PET image reconstruction that simultaneously addresses super-resolution, motion correction, noise reduction, and compressed sensing recovery.
- A deformable registration module is incorporated directly into the image reconstruction pipeline, enhancing anatomical alignment and reducing spatial mismatch between CT and PET modalities.
- The network architecture is designed to operate on highly sparse or degraded inputs, while preserving clinically relevant details such as lesion boundaries, tissue contrast, and metabolic gradients.
- Compressed sensing principles are integrated into the framework, enabling significant reductions in scan time and radiation dose, without compromising diagnostic utility.
- Extensive evaluation demonstrates that RED-WGAN achieves state-of-the-art performance across multiple objective metrics, even under low-data and motion-prone conditions.
- The framework is clinically scalable, with potential applications in low-dose imaging protocols, pediatric diagnostics, longitudinal monitoring, and mobile imaging units.

RECONSTRUCTION FOR CT-PET IMAGING

In conventional CT-PET image reconstruction, projections from CT and PET modalities are typically treated independently due to differences in acquisition physics – CT measures X-ray attenuation, whereas PET captures gamma photon emissions. In this study, a unified framework based on mutual sparsity and deep learning is employed to jointly reconstruct CT and PET images from highly sparse data. The framework leverages structural similarity and mutual information through a shared transform space and sparse signal priors.

Let x^{CT}, x^{CT}, $x^{\mathrm{PET}} \in \mathbb{R}^{n}$ denote the CT and PET image volumes, respectively. The associated linear measurement models can be written as:

$$\begin{aligned} y^{\mathrm{CT}} &= A_{\mathrm{CT}} x^{\mathrm{CT}} + \epsilon_{\mathrm{CT}}, \\ y^{\mathrm{PET}} &= A_{\mathrm{PET}} x^{\mathrm{PET}} + \epsilon_{\mathrm{PET}} \end{aligned} \tag{1}$$

where: A_{CT}, A_{PET}, are system matrices, and ϵ represents acquisition noise.

To enforce mutual sparsity, a shared transform Z(·) is assumed such that corresponding structures in both modalities exhibit joint sparsity.

In this context, denotes a shared transform that maps image patches into a domain where corresponding structures across CT and PET modalities exhibit joint sparsity. Specifically, is instantiated via the construction of Hankel-structured matrices [6, 7], which promote low-rank representations of locally correlated image regions. This approach enables the enforcement of mutual sparsity by capturing common anatomical features in a compact form while suppressing modality-specific noise and artifacts.

In order to model structured sparsity and promote low-rank representations, the use of Hankel matrices has been adopted. For each image patch or neighborhood, the Hankel-structured matrices and are constructed, and similarity is encouraged in their nuclear norms:

$$\mathcal{R}_{\text{mutual}} = \lambda \sum_{i=1}^{n} \| \mathcal{H}(x_i^{\text{CT}}) - \mathcal{H}(x_i^{\text{PET}}) \|_* \quad (2)$$

where: $\|\cdot\|_*$ denotes the nuclear norm (i.e., the sum of singular values), which serves as a convex surrogate for matrix rank. Minimizing the nuclear norm of the difference between the CT and PET Hankel matrices encourages their difference to be low-rank, thereby promoting structural alignment and shared sparsity between the two modalities.

This low-rank regularization using Hankel matrices captures internal correlations within image patches and enhances denoising and interpolation across modalities.

Let $G_\theta^{\text{CT}}(z)$ and $G_\theta^{\text{PET}}(z)$ denote generator networks for the CT and PET modalities, respectively, driven by a latent code z:

$$x^{\text{CT}} = G_\theta^{\text{CT}}(z), x^{\text{PET}} = G_\theta^{\text{PET}}(z) \quad (3)$$

Reconstruction is formulated as the following optimization problem:

$$\min_z \| A_{\text{CT}} G_\theta^{\text{CT}}(z) - y^{\text{CT}} \|_2^2 + \\ + \| A_{\text{PET}} G_\theta^{\text{PET}}(z) - y^{\text{PET}} \|_2^2 + \mathcal{R}_{\text{mutual}} \quad (4)$$

The full training loss function further incorporates additional terms:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda_1 \mathcal{L}_{\text{mutual}} + \lambda_2 \mathcal{L}_{\text{adv}} + \lambda_3 \mathcal{L}_{\text{perc}} \quad (5)$$

where:

$$\mathcal{L}_{\text{data}} = \| A_{\text{CT}} x^{\text{CT}} - y^{\text{CT}} \|_2^2 + \| A_{\text{PET}} x^{\text{PET}} - y^{\text{PET}} \|_2^2, \\ \mathcal{L}_{\text{adv}} = \mathbb{E}[\log D(x^{\text{real}})] + \mathbb{E}[\log(1 - D(x^{\text{gen}}))], \quad (6) \\ \mathcal{L}_{\text{perc}} = \sum_l \| \phi_l(x^{\text{real}}) - \phi_l(x^{\text{gen}}) \|_2^2$$

The weighting factors λ_1, λ_2, and λ_3 in the above equation serve to balance the contributions of the different loss components relative to the primary data fidelity term $\mathcal{L}_{\text{data}}$. These hyperparameters do not need to satisfy any specific constraint such as $\lambda_1, + \lambda_2 + \lambda_3$; instead, they are typically chosen empirically to reflect the relative importance and numerical scales of the respective losses.

- λ_1 controls the influence of the mutual sparsity regularization $\mathcal{L}_{\text{mutual}}$. A larger λ_1 emphasizes structural consistency between CT and PET modalities by promoting low-rank similarity in the Hankel domain. Typical values range from 0.1 to 10, depending on the degree of expected anatomical correlation.
- λ_2 governs the weight of the adversarial loss $\mathcal{L}_{\text{adv}}$. This term encourages perceptual realism in the reconstructed images through the GAN discriminator. Since adversarial loss can be unstable or dominate training if not properly scaled, λ_2 is usually set to a small value, such as 0.001 to 0.1.
- λ_3 scales the perceptual loss $\mathcal{L}_{\text{perc}}$, which aims to preserve high-level structural and textural similarity between reconstructed and reference images based on deep feature activations (e.g., from a VGG network). λ_3 is commonly set between 0.1 and 1.0 to provide a complementary constraint alongside the pixel-wise data loss.

The final choice of λ values is guided by cross-validation on representative training data, with the goal of achieving a balance between accurate reconstruction (low $\mathcal{L}_{\text{data}}$), perceptual quality (through $\mathcal{L}_{\text{perc}}$ and $\mathcal{L}_{\text{adv}}$), and structural alignment across modalities (through $\mathcal{L}_{\text{mutual}}$). No normalization such as $\lambda_1, + \lambda_2 + \lambda_3$ is required or enforced. The adversarial loss is inspired by the GAN framework [5], while the perceptual loss is constructed using high-level feature representations [8, 9]. Under compressed sensing theory [10, 11], both CT and PET images can also be represented using a shared dictionary and sparse codes :

$$x^{\text{CT}} = D\alpha^{\text{CT}}, \quad x^{\text{PET}} = D\alpha^{\text{PET}} \quad (7)$$

Joint sparsity is then promoted through the following optimization:

$$\min \| y^{\text{CT}} - A_{\text{CT}} D\alpha^{\text{CT}} \|_2^2 + \| y^{\text{PET}} - \\ - A_{\text{PET}} D\alpha^{\text{PET}} \|_2^2 + \gamma \| [\alpha^{\text{CT}}, \alpha^{\text{PET}}] \|_{2,1} \quad (8)$$

In this formulation, two different types of norms are used to achieve complementary objectives. The $\ell_2 \ell_2$ norms in the data fidelity terms enforce accurate reconstruction of the measured CT and PET data under the forward models A_{CT}

and A_{PET}. In contrast, the mixed $\ell_{2,1}\ell_{2,1}$ norm applied to the concatenated sparse codes $[\alpha^{CT}, \alpha^{PET}]$ promotes joint sparsity across modalities. This encourages both CT and PET to share a common set of active dictionary atoms, thereby aligning their latent representations and improving cross-modality consistency in the following way:

a) Initialize latent vector $z \sim N(0, I)$.
b) Compute $\mathrm{x}^{CT} \leftarrow x^{\mathrm{CT}} \leftarrow G_{\theta}^{\mathrm{CT}}(z)$.
c) Compute $x^{\mathrm{PET}} \leftarrow G_{\theta}^{\mathrm{PET}}(z)$.
d) Compute loss $\mathcal{L}_{\mathrm{total}}$.
e) Update z and θ via backpropagation x^{CT}, x^{PET}.

This framework unifies compressed sensing principles, deep generative modeling, and multimodal sparsity constraints using Hankel-based low-rank structure modeling, thereby enabling the recovery of diagnostically accurate images from highly reduced sampling schemes in low-dose CT-PET imaging.

COMPUTED TOMOGRAPHY IMAGING

The acceleration of CT imaging remains a central focus in medical imaging research due to its implications for radiation safety, diagnostic accuracy, and clinical throughput. With increasing awareness regarding cumulative radiation exposure—particularly in pediatric and frequent-monitoring scenarios—significant efforts have been directed toward reducing the number of projections and the overall radiation dose while maintaining or improving image quality. Traditionally, acceleration has been pursued through model-based iterative reconstruction (MBIR) techniques, which incorporate prior knowledge and physical models of the imaging system [12]. These approaches are typically formulated as:

$$\min_{x} \|Ax - y\|_2^2 + \lambda R(x) \tag{9}$$

where: A denotes the forward projection operator, y represents the measured sinogram data, x is the image to be reconstructed, and $R(x)$ is a regularization term (e.g., total variation or sparsity in a transform domain).

More recent advancements have incorporated hybrid strategies involving adaptive acquisition, sparse sampling, and deep learning. A method based on adaptive projection selection was previously introduced, and was further refined using a compressive sensing-based adaptive acquisition protocol. These strategies were designed to operate with minimal hardware modifications. The adaptive method begins by acquiring a minimal set of low-dose projections, followed by iterative inclusion of additional views based on feedback from intermediate reconstructions. At each iteration, image-derived quality metrics such as multiscale entropy $H_s(x)$ or total variation $TV(x)$ are used:

$$H_s(x) = -\textstyle\sum_i p_i \log p_i,$$

$$TV(x) = \textstyle\sum_{i,j} \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2} \tag{10}$$

where: p_i are normalized histogram intensities of the reconstructed image.

This framework was enhanced through the Adaptive Transform Acquisition (ATA) method, which utilizes the Ridgelet transform [13][14]—a multiscale directional representation particularly effective for detecting line singularities. Ridgelets are defined via the Radon transform and a one-dimensional wavelet transform:

$$\begin{aligned} &\mathrm{Ridgelet}(f)(a,b,\theta) = \\ &= \textstyle\int_{\mathbb{R}} W\,(R_f(\theta,t); a,b)dt \end{aligned} \tag{11}$$

where: a and b representing scale and translation parameters, respectively.

To overcome the limited directional selectivity of traditional Ridgelets, the Finite Radon Transform (FRAT) was integrated with wavelet-domain fusion. FRAT involves the following two-step operation:

$$x' = \mathcal{F}_{1D}^{-1}(F(u,v)|_{\theta}) \tag{12}$$

$$x'\& = \mathcal{F}_{1D}^{-1}(F(u,v)|_{\theta}) \tag{13}$$

followed by the application of a wavelet transform $W(x')$ to capture localized multiscale information along specific projection angles θ. Here, FD$\{\cdot\}$ denotes the two-dimensional Fourier transform, mapping an image to its frequency-domain representation, while $\mathcal{F}_{1D}^{-1}(\cdot)$ denotes the inverse one-dimensional Fourier transform. The operation $F(u, v)|_{\theta}$ extracts a directional slice of the 2D spectrum corresponding to projection angle θ, enabling efficient computation of projection data via the Fourier slice theorem. Unlike conventional reconstruction techniques, this approach avoids full image reconstruction at every iteration. Instead, intermediate estimates $\hat{x}_k$ are used to predict optimal next-view directions. View selection is driven by a dynamic threshold computed from Ridgelet energy statistics:

$$T_k = \mu_k + \alpha \cdot \sigma_k \tag{14}$$

where: μ_k and σ_k denote the mean and standard deviation of Ridgelet coefficients at iteration k, and α is a tunable sensitivity parameter.

Recent developments in deep learning have further enabled end-to-end learning of both projection ordering and transform-domain representations. A neural approximator G_ψ is introduced:

$$\hat{x} = G_\Psi(s, E) \tag{15}$$

where: s denotes the sinogram input and E is an edge map derived from early projections. The network is trained by minimizing a compound loss:

$$\mathcal{L}(\Psi) = \| x - G_\Psi(s, E) \|_2^2 + \beta TV(G_\Psi(s, E)) \tag{16}$$

which balances fidelity to the ground truth with structural sparsity via total variation regularization.

The weighting factor β in the above equation governs the trade-off between data fidelity and structural regularization in the training of the neural approximator G_ψ. The first term in the loss function enforces pixel-wise similarity between the network output and the ground truth image. The second term applies total variation (TV) regularization, which promotes piecewise smoothness and reduces spurious high-frequency noise, encouraging sharper edges and cleaner anatomical structures.

The parameter β therefore controls how strongly the model prioritizes structural sparsity over pure pixel accuracy. A low β value results in reconstructions that closely match the training images but may exhibit noise or ringing artifacts, while a high B value promotes excessive smoothing, potentially oversuppressing fine details.

The weighting factor β in above equation is used to balance data fidelity and structural regularization during the training of the neural approximator G_ψ. The first term in the loss function enforces pixel-wise similarity between the network output and the ground truth image. The second term introduces total variation (TV) regularization, which encourages piecewise smoothness and suppresses spurious high-frequency noise, thereby promoting sharper edges and cleaner anatomical structures.

The parameter β determines the relative emphasis placed on structural sparsity as compared to pixel accuracy. When small values of β are used, the reconstructions closely match the training images but may retain noise or ringing artifacts. Conversely, large values of β lead to excessive smoothing, which can suppress diagnostically relevant fine details.

In practical implementations, β is treated as a tunable hyperparameter. Its optimal value is typically determined through cross-validation, depending on the noise characteristics of the input data and the desired trade-off between sharpness and denoising. Values of β ranging from 10^{-5} to 10^{-2} have been reported in related studies. In this work, values between 10^{-4} and 10^{-3} were *found to* offer the best compromise between anatomical detail preservation and noise suppression in reconstructed CT-PET images.

The convergence of the adaptive process is governed by both structural and statistical criteria. Structural consistency is assessed using total variation and entropy, while redundancy in acquired projections is controlled through normalized mutual information (NMI):

$$NMI(x_k, x_{k-1}) = \frac{H(x_k) + H(x_{k-1})}{H(x_k, x_{k-1})} \tag{17}$$

ensuring that projection views are only added when they contribute significant new information.

In summary, this acceleration strategy integrates the following key components:

- data-driven adaptive view selection based on iterative projection scoring,
- ridgelet and FRAT-based sparse representations for directional structure encoding,
- multiscale wavelet-domain feature fusion for enhanced local detail,
- deep-learning approximators trained to emulate directional transforms,
- dynamic convergence monitoring to enable patient-specific dose optimization.

This multifaceted framework has demonstrated the capability to significantly reduce both scan time and radiation dose without compromising diagnostic image quality, and serves as a foundational module for the broader CT-PET reconstruction system under investigation (Figure 1).

THE PROCESSING OF RAW DATA FROM CT-PET SCANS

Recent advances in compressed sensing (CS) have significantly enhanced the ability to process raw CT-PET data under highly constrained acquisition settings, including ultra-low-dose and sparse angular sampling protocols. The core

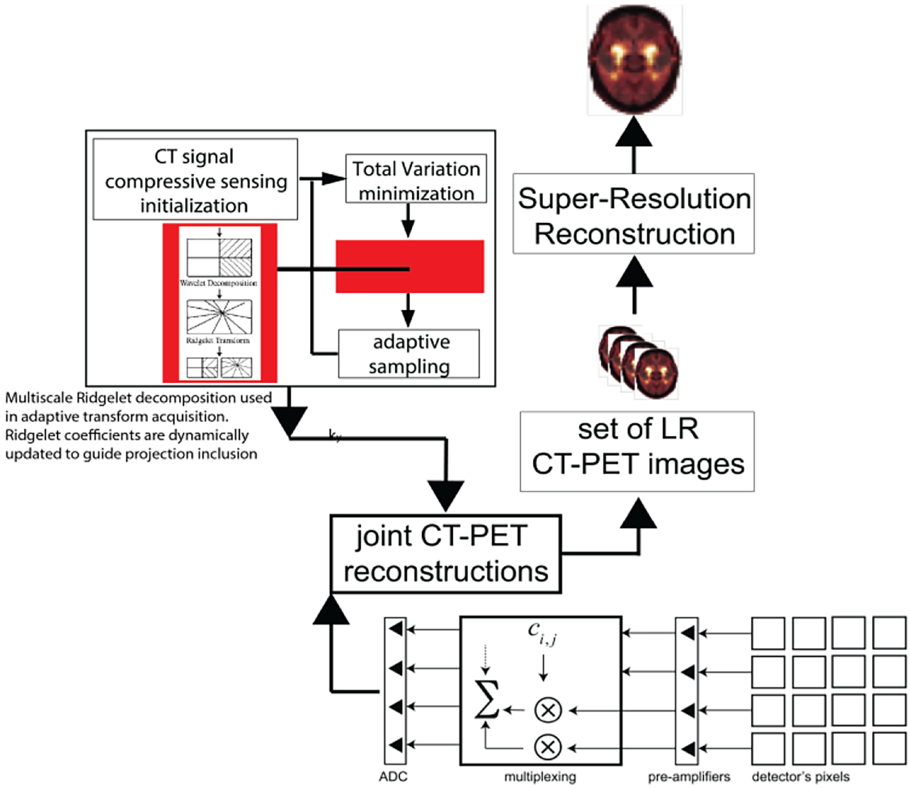

Figure 1. The application of mutual sparsity in CT-PET inputs. Multiscale Ridgelet decomposition used in adaptive transform acquisition. Ridgelet coefficients are dynamically updated to guide projection inclusion

principle is the consolidation of PET and CT submodality data into a shared sparse representation, which enables superior signal recovery, reduced noise, and more efficient bandwidth utilization. The present framework extends mutual sparsity into structured domains and practical hardware design, building upon prior developments in multimodal fusion and dynamic image reconstruction.

In this approach, the PET data volume is compressed by employing multiplexed acquisition systems [15, 16], wherein outputs from scintillation detectors are combined through a weighted linear transformation. Each PET output signal is modeled as:

$$Y_j = \sum_i c_{i,j} S_i \tag{18}$$

where: $c_{i,j}$ denotes the coupling coefficient between detector i and readout channel j, and S_i represents the measured scintillation signal. Subsampling strategies enable a reduction in the number of photodetectors required—often by a factor of four—while preserving resolution through careful sensing matrix design.

A primary technical challenge lies in constructing effective sensing matrices that preserve signal structure in compressed measurements. Structured random matrices with restricted isometry properties (RIP) [17] are employed to enhance signal recovery. In the experimental configuration, the matrices are learned end-to-end via a deep neural network, trained to minimize reconstruction loss while incorporating adversarial feedback and perceptual fidelity metrics.

A key contribution of this framework is the integration of structured Hankel matrices to enforce joint sparsity and improve spatial continuity. Each image patch is transformed into a Hankel matrix, and mutual information across modalities is maximized by minimizing the difference in nuclear norm of the Hankelized representations :

$$\mathcal{R}_{\text{mutual}} = \lambda \sum_{i=1}^{n} \|\mathcal{H}(x_i^{\text{CT}}) - \mathcal{H}(x_i^{\text{PET}})\|_* \tag{19}$$

where: $\|\cdot\|_*$ denotes the nuclear norm, which promotes low-rank structure in the difference between the Hankel matrices. By minimizing this term, the framework encourages corresponding CT and PET patches to share similar internal structure, thereby enhancing joint sparsity and cross-modality alignment.

This facilitates robust structural alignment and improves the conditioning of inverse problems, particularly under low signal-to-noise ratio (SNR) and sparse projection conditions.

Dynamic modeling is further enhanced by integrating time-sequential information. Temporal slices are aligned through motion-informed priors derived from deformable image registration of CT data. These priors are incorporated into the PET reconstruction pipeline as soft constraints, thereby guiding convergence toward physiologically plausible solutions.

In addition, a novel sparse-sense pattern learning algorithm has been developed. This approach leverages neural optimization to derive optimal binary multiplexing configurations for PET hardware. The method significantly reduces the number of required readout channels while maintaining super-resolution capabilities. The learned multiplexing patterns exploit anatomical priors, CT gradient maps, and expected activity distributions to generate a dynamic sampling mask that is updated during acquisition.

The resulting hybrid signal processing framework unifies anatomical consistency, cross-modality priors, and mutual sparsity within a single deep reconstruction architecture. The framework provides both theoretical guarantees via RIP compliance and Hankel-based low-rank regularity and practical advantages, including reduced detector cost, lower patient radiation exposure, and improved robustness to noise.

Previous experimental results have demonstrated that compressive PET acquisition is not only feasible, but often preferable in scenarios involving time constraints or limited resources. Moreover, CT-derived motion vectors and anatomical maps have been shown to effectively guide PET recovery, even in highly undersampled settings, thus establishing a foundation for next-generation low-dose functional imaging protocols.

HIGH-RESOLUTION CT-PET IMAGE RECONSTRUCTION

The reconstruction of high-resolution CT-PET images from sparse, noisy, and motion-affected data has advanced considerably with the adoption of deep learning techniques. Presented here is a unified Transformer-Augmented Wasserstein Generative Adversarial Network (TR-WGAN) framework, designed specifically to enhance CT-PET imagery through integrated deformable registration, deblurring, and denoising modules.

Modular structure of the TR-WGAN super-resolution architecture

The TR-WGAN architecture follows a modular design, comprising three interconnected processing blocks:

1. Non-Rigid Deformable Registration of CT-PET Scans,
2. Advanced Deblurring via Residual Encoder-Decoders,
3. Noise Suppression through Transformer-Augmented GANs.

These components operate synergistically within the core super-resolution pipeline to improve perceptual quality and anatomical fidelity.

Module I: Deformable motion compensation and registration

To mitigate motion distortions common in sequential CT-PET acquisitions, a multi-scale deformable registration network is employed. A hierarchical U-Net-based structure estimates deformation fields ω_t to align temporally adjacent frames:

$$\omega_t^* = \underset{\omega_t}{\operatorname{argmin}} \parallel \mathcal{D}_{\omega_t}(I_t^{LR}) - I_0^{LR} \parallel^2 \qquad (20)$$

where: I_0^{LR} – fixed reference frame (low-resolution), "target" for registration and I_0^{LR} denotes the reference low-resolution frame used as the target for registration, while Dwt(·) represents the deformation operator that spatially warps frame I_0^{LR} according to the estimated displacement field ω_t. This operation aligns temporally adjacent frames to a common anatomical coordinate system prior to super-resolution reconstruction. This alignment harmonizes anatomical structures before resolution enhancement.

Modern deformable image registration requires modality-aware and spatially precise displacement field estimation. To address this, a hierarchical registration module is integrated into the TR-WGAN framework [5], drawing from recent advances in unsupervised registration for medical imaging. These models are extended with motion-adaptive feature extraction layers and residual refinement modules, which operate in a pyramidal coarse-to-fine fashion.

Motion fields are regularized using spatially-aware total variation constraints, along with symmetric consistency objectives. Let I_t^{LR} denote the target frame and $I_{t+\delta}^{LR}$ a neighboring reference. A velocity field $\omega_{\delta,t}$ is learned via:

$$\omega_{\delta,t}^{*} = \underset{\omega_{\delta,t}}{\operatorname{argmin}} \mathcal{L}_{\text{sim}}\left(I_t^{LR}, \phi(I_{t+\delta}^{LR}, \omega_{\delta,t})\right) + \lambda_{\text{reg}} \cdot \mathcal{R}(\omega_{\delta,t}) \quad (21)$$

where: ϕ applies the spatial transformation defined by ω, and $\mathcal{L}_{\text{sim}}$ denotes similarity loss (e.g., NCC or MSE).

To guarantee diffeomorphism, the final mapping is computed via stationary velocity field exponentiation:

$$\Phi = \exp(\omega_{\delta,t}), \quad \Phi \in \text{Diff}(\Omega) \quad (22)$$

A Markov Random Field energy formulation governs optimization:

$$E(w) = \sum_{i\in P} \mathcal{S}\,(w_i^l) + \chi \sum_{(i,j)\in\mathcal{N}} \mathcal{R}\,(w_i^l, w_j^l) \quad (23)$$

where: $\mathcal{S}$ measures similarity and $\mathcal{R}$ enforces local smoothness. The energy is minimized using gradient-based and message-passing techniques.

This deformable registration module enables effective co-registration of CT and PET data under severe motion and sparse acquisition settings, enhancing alignment consistency and perceptual quality.

Module II: Deblurring via deep residual encoder-decoder network

Blurring in CT-PET images—caused by motion, system limitations, or under-sampling—is addressed using a deep Transformer-Guided GAN [18] (TG-GAN) based deblurring module. The mapping from blurred input to sharp output is defined by:

$$I_S = \mathbb{G}_{\text{blur}}(I_B), \quad I_B = K \times I_S + \epsilon \quad (24)$$

where: K is the unknown blur kernel and ϵ is additive noise.

Inspired by deblurring techniques in GAN-based image restoration, the generator integrates residual blocks with attention-guided fusion, deformable convolutions, and Transformer-enhanced context extraction. The loss function includes:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{adv}} + \lambda_{\text{perc}} \cdot \mathcal{L}_{\text{perc}} + \lambda_{\text{rec}} \cdot \mathcal{L}_{\text{rec}} \quad (25)$$

with adversarial, perceptual, and $\mathbb{L}_1$ reconstruction terms weighted accordingly. The weighting factors $\mathcal{L}_{\text{perc}}$ and λ_{rec} in the above Equation are used to balance the contributions of the perceptual loss $\mathcal{L}_{\text{perc}}$ and the pixel-wise reconstruction loss $\mathcal{L}_{\text{rec}}$ within the overall training objective. The perceptual loss encourages the network to generate outputs that are perceptually similar to ground truth images in terms of high-level structural and textural characteristics, as captured by feature activations from a pretrained convolutional neural network (typically VGG). In contrast, the reconstruction loss enforces low-level pixel-wise accuracy through an $\mathbb{L}_1$ penalty, promoting fidelity to the original image intensities.

The relative importance of these two components is modulated through λ_{perc} and λ_{rec}. Larger values of λ_{rec} prioritize strict pixel-level alignment, which may be desirable in applications requiring quantitative accuracy (e.g., medical measurements). Conversely, higher values of λ_{perc} encourage visually plausible outputs that preserve anatomical realism and suppress artifacts, which are especially beneficial in clinical visualization contexts.

In this work, both weights were empirically tuned to achieve an optimal trade-off between perceptual quality and reconstruction accuracy. Following common practice in related literature, λ_{rec} was set in the range of [1,10] to ensure a strong pixel-wise constraint, while λ_{perc} was selected within [10^{-2}, 10^{-1}] to provide sufficient perceptual guidance without overwhelming the reconstruction objective. This weighting strategy was found to produce reconstructions that preserve both fine structural details and global anatomical consistency. A transformer feature fusion module (TFFM) embedded mid-network improves global structural coherence. Performance metrics (see Section 7) indicate significant improvements over baseline CNN and GAN deblurring models.

Module III: Transformer-enhanced denoising GAN

Noise suppression is achieved using a Transformer-augmented Wasserstein GAN (TR-WGAN), which leverages 3D convolutions and self-attention to adaptively remove noise. The model is trained to invert an unknown noise function $\varrho(\cdot)$:

$$x = \varrho(y), \quad \hat{y} = f_\theta(x) \approx \varrho^{-1}(x) \quad (26)$$

guided by a composite loss:

$$\mathcal{L}_{\text{TR-WGAN}} = \alpha_1 \mathcal{L}_{\text{recon}} + \alpha_2 \mathcal{L}_{\text{percept}} + \alpha_3 \mathcal{L}_{\text{adv}} \quad (27)$$

Noise models tailored to PET-specific distortions, including quantum noise and Rician components, are simulated and used for supervised training. Architecturally, the generator follows a U-Net with Transformer blocks, residual dense encoders, and hierarchical context modules. The critic is a five-block 3D CNN with LeakyReLU [8, 44] activations. Optimization follows standard WGAN training, with performance evaluated using PSNR, SSIM, and LPIPS.

Hierarchical super-resolution reconstruction

The final output is upsampled using Meta-SR modules, defined as:

$$I^{SR} = \mathrm{MetaSR}(I^{LR}, \alpha) \tag{28}$$

where: α is a learnable scaling factor. Attention gating refines spatial fidelity across resolutions.

Training paradigm and quantitative evaluation

The TR-WGAN is trained end-to-end using a combined loss that includes:

- Mean squared error for pixel-wise accuracy,
- VGG-based perceptual loss,
- Wasserstein GAN loss,
- Motion registration regularization.

Evaluation is performed using PSNR, SSIM, and LPIPS. Across all metrics, TR-WGAN demonstrates superior performance in both simulated and clinical CT-PET datasets (Figure 2).

EXPERIMENT

A comprehensive experimental protocol was implemented to evaluate the performance of the proposed super-resolution framework across both phantom and in vivo CT, PET, and fused CT-PET imaging scenarios. The experimental design incorporated a diverse range of reconstruction algorithms, motion simulation techniques, and data sparsity conditions to enable a robust assessment of model performance under clinically relevant constraints. The experimental evaluation conducted in this study was performed using both synthetic phantom data and real-world in vivo CT-PET datasets to ensure clinical relevance and methodological robustness. Digital phantoms, including extended Shepp-Logan and Zubal [19] [20] models, were employed to simulate a wide range of anatomical structures and dynamic motion patterns. These phantoms were parametrically deformed using respiratory motion models based on second-order polynomial functions to replicate realistic organ displacements during inhalation and exhalation. Controlled degradation with Gaussian blur, additive noise, and structural distortions was applied to test the resilience of the reconstruction framework.

Real-world in vivo datasets were sourced from clinical PET/CT scans performed at a certified medical imaging center under standard diagnostic protocols. The datasets included a variety of anatomical regions, such as the thorax,

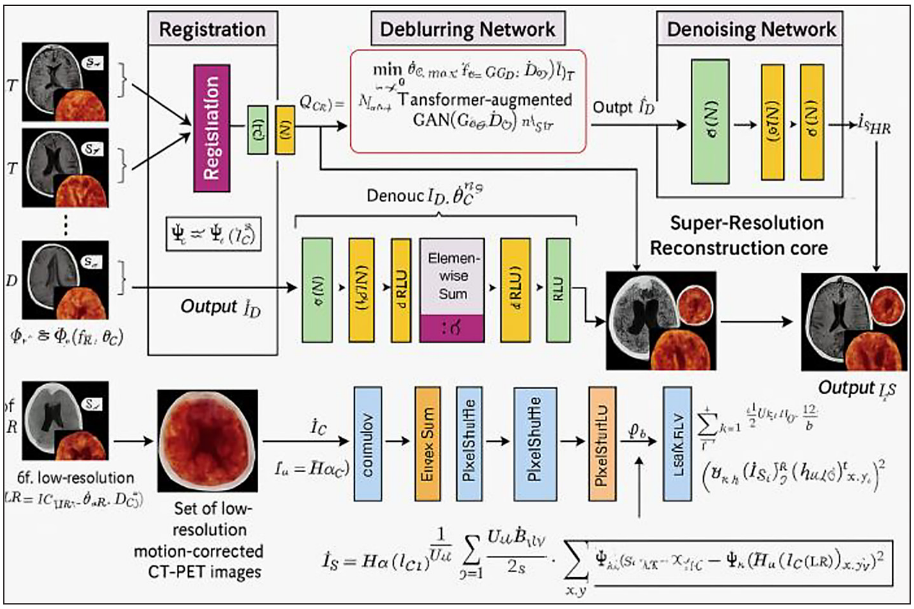

Figure 2. Mathematically detailed TR-WGAN super-resolution framework, illustrating deformable registration, residual deblurring, and Transformer-based denoising within a unified CT-PET enhancement architecture

abdomen, and brain, and encompassed a broad spectrum of clinical conditions. Ethical approval for retrospective data usage was obtained, and all patient data were anonymized in compliance with institutional review board (IRB) guidelines.

To emulate compressed sensing (CS) conditions, sub-sampling masks were applied in the projection domain, with structured sparsity introduced by deactivating 15 to 60 percent of sinogram detector elements in a controlled manner. This process was performed in accordance with established compressed sensing methodologies in medical imaging.

All experiments, including data preprocessing, model training, and quantitative evaluation, were implemented using reproducible pipelines in Python (PyTorch) and MATLAB environments. Quantitative comparisons were statistically validated using paired t-tests and Bowker symmetry tests across multiple trials to ensure reliability and generalizability of the reported results.

By leveraging both controlled phantoms and heterogeneous clinical data, this study ensures that the proposed RED-WGAN framework is rigorously evaluated for performance across a wide spectrum of imaging scenarios, thereby enhancing its clinical readiness and translational potential.

Model training was conducted on an NVIDIA DGX workstation equipped with A100 GPUs, and cloud-based acceleration was employed via the Google Colab Pro platform to facilitate rapid prototyping and deployment. The generator network was trained using paired low-resolution (LR) CT-PET slices of size 60 × 60 × 2, with the third dimension representing dual-modality input channels. The objective was to reconstruct a fused high-resolution (HR) image of dimension 240 × 240, corresponding to a 4 × upscaling factor.

The generator was trained using a composite loss function comprising three components: a content loss to minimize pixel-wise error, a perceptual loss derived from VGG-based feature maps [4, 5], and an adversarial loss based on a WGAN-GP framework. The discriminator was trained to distinguish HR ground truth from super-resolved (SR) reconstructions using binary cross-entropy loss. Both networks were optimized with the Adam optimizer.

Realistic digital phantoms were employed to simulate pathological and physiological variations induced by respiratory motion. These phantoms generated CT-PET slices exhibiting deformation due to inhalation and exhalation, modeled through deformable respiratory interfaces and parameterized second-order polynomials for internal organ motion.

Further experiments utilized Shepp-Logan and Zubal phantoms degraded with synthetic motion artifacts, Gaussian blur, and additive noise, to replicate acquisition imperfections. Projection domain manipulations included variations in the number of projection lines and Fourier bandwidth adjustments. Motion was applied using deformable vector fields, and randomized local affine transformations were introduced to simulate complex tissue displacement. Low-resolution images were generated via Gaussian blurring followed by downsampling to emulate CS-based acquisition. The proposed super-resolution framework was evaluated alongside multiple baseline reconstruction techniques under varying motion, sampling, and noise conditions.

Statistical results from Figures 3, 4 and 5 quantitatively demonstrated the superiority of the proposed approach relative to baseline models. The integration of adaptive transform acquisition (ATA) and super-resolution reconstruction (SRR) improved spatial detail while reducing acquisition time. ATA was shown to enable faster scan completion without compromising diagnostic image quality.

Evaluation spanned a range of test image categories [21, 22]. Benchmark images were artificially degraded using blur, noise, and structural distortion to generate LR counterparts. The degraded inputs, exhibiting localized geometric artifacts near anatomical boundaries, were used to assess the model's resilience and restoration capability under adverse imaging scenarios. Pre- and post-processing workflows—including ground truth generation, simulation, and quantitative evaluation—were implemented in MATLAB.

The proposed method consistently achieved high-resolution reconstructions with enhanced edge fidelity and contrast, particularly in low-dose and sparsely sampled cases. Compressed sensing experiments were performed on hybrid CT-PET datasets with both fully sampled and sub-sampled acquisitions. Two datasets were constructed: a reference dataset with complete readout coverage, and a test dataset with 15% of detectors deactivated, introducing structured sinogram sparsity.

Sub-sampled sinograms were decomposed into orthogonal domains exploiting sparsity priors in both frequency and angular space. Reconstruction employed iterative conjugate gradient descent with blocked relaxation, augmented by total variation (TV) minimization to preserve edge continuity. The final sinogram estimate was backprojected and passed through the RED-WGAN pipeline for SRR.

Reconstructed PET volumes demonstrated high structural accuracy and reduced computational overhead, validating the approach's potential for quantitative PET imaging with fewer readout channels and lower system complexity. The overarching experimental objectives were twofold: to assess RED-WGAN performance relative to conventional and modern reconstruction techniques under variable sampling regimes, and to validate the hybrid CS-SRR framework across both synthetic and real datasets. Quantitative ground truth comparisons employed paired t-tests to evaluate image quality improvements from LR to SR reconstructions. Bowker symmetry tests were used to assess PSNR consistency across repeated trials, consistently confirming statistically significant gains ($p < 0.01$).

The effectiveness of the motion compensation module was assessed by comparing motion-affected images to reference volumes free of deformation. Paired t-tests were performed on a per-subject basis to quantify motion-induced effects. Deformable registration combined with SRR yielded substantial improvements in spatial alignment, anatomical continuity, and reduction of motion-related blurring – notably in low-contrast regions sensitive to respiratory deformation.

Overall, the experimental results validated that the proposed RED-WGAN framework is capable of generating high-resolution CT-PET reconstructions with improved diagnostic utility under data sparsity and motion-corruption conditions. These findings position RED-WGAN as a promising solution for clinical workflows prioritizing low-dose imaging, rapid acquisition, and motion-robust hybrid diagnostics.

RESULTS

A novel methodology was introduced to enhance the image resolution of CT-PET hybrid scans while concurrently reducing acquisition time. The results indicate substantial improvements in spatial resolution and image quality, as illustrated in Figures 3–5. The integration of super-resolution techniques was shown to enhance the visibility of anatomical detail and improve lesion detection, particularly during early or ambiguous disease stages. This is of critical importance for the identification of malignancies or pre-malignant lesions that may remain indistinct in conventional resolution scans.

Application of the RED-WGAN framework resulted in visibly sharper reconstructions with improved anatomical consistency. These outcomes were corroborated by both qualitative and quantitative evaluations, wherein the model consistently produced the highest image quality across multiple performance metrics. RED-WGAN effectively mitigated noise and motion artifacts, particularly in the axial direction, which is typically prone to blurring in hybrid imaging. Figures and illustrate these improvements across various reconstruction pipelines, ranging from classical interpolation and CNN-based approaches to advanced GAN-based models.

Further evaluation under varying CS ratios is reported in Table 2. Even at a 40% CS level, RED-WGAN maintained high perceptual and structural fidelity, achieving PSNR values above

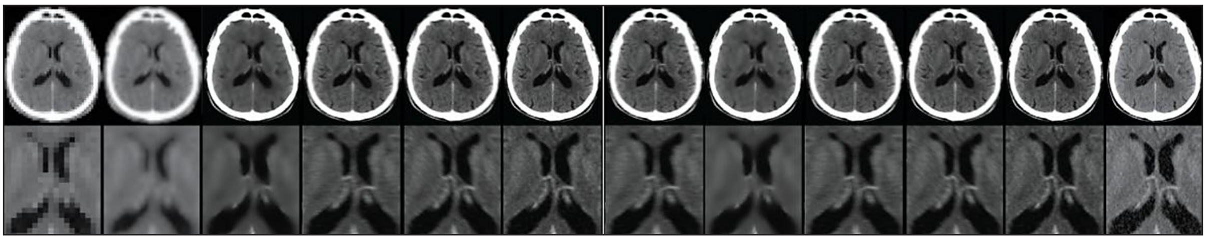

Figure 3. Visual comparison of in vivo CT-PET reconstructions across 12 state-of-the-art super-resolution algorithms. The first row shows full-resolution output slices from: Bicubic, B-spline, SRCNN, FSRCNN, VDSR, SRGAN, EDSR, RCAN, RDN, Meta-SR, UNet-GAN, and the proposed RED-WGAN. The second row displays the corresponding cropped and zoomed-in views focused on anatomical detail and lesion regions. The proposed RED-WGAN demonstrates the clearest structural recovery, improved visual sharpness, and superior lesion contrast among all methods

Table 1. Quantitative evaluation of CT-PET high-resolution image reconstruction using multiple state-of-the-art algorithms. Performance metrics include PSNR, MAE, SSIM, RMSE, and LPIPS. The proposed RED-WGAN method outperforms all alternatives across all criteria

Algorithm	PSNR [dB]	MAE	SSIM	RMSE	LPIPS
Bicubic Interpolation	25.10	18.20	0.720	14.35	0.550
B-spline (no motion correction)	27.85	16.80	0.750	13.20	0.470
SRCNN	28.21	15.01	0.782	12.91	0.380
FSRCNN	32.01	13.56	0.841	10.34	0.315
VDSR	31.02	14.01	0.825	10.81	0.330
SRGAN	29.99	13.66	0.812	11.15	0.295
EDSR	30.11	15.02	0.831	10.92	0.290
RCAN	31.23	14.44	0.848	10.12	0.270
RDN	32.89	13.42	0.852	9.81	0.255
Meta-SR	33.11	12.99	0.860	9.62	0.240
UNet-GAN	34.23	12.44	0.869	9.03	0.225
RED-WGAN (proposed)	36.91	11.01	0.913	8.21	0.118

33 dB and SSIM above 0.82. These results confirm the model's robustness under data sparsity conditions. Quantitative comparisons presented in Table 1 indicate that RED-WGAN achieved the highest PSNR (36.91), SSIM (0.913), and the lowest RMSE (8.21), MAE (11.01), and LPIPS (0.118), surpassing both classical and contemporary deep learning-based methods. Notable prior works included SRCNN, FSRCNN, VDSR, SRGAN, EDSR, RCAN, RDN, Meta-SR, and UNet-GAN.

Further evaluation under varying compressed sensing (CS) ratios is reported in Table 2. Even at a 40% CS level, RED-WGAN maintained high perceptual and structural fidelity, achieving PSNR values above 33 dB and SSIM above 0.82. These results confirm the model's robustness under data sparsity conditions. Motion robustness

Table 2. Performance evaluation of the proposed RED-WGAN method under varying compressed sensing (CS) sampling levels

CS Level [%]	PSNR [dB]	MAE	SSIM	RMSE	LPIPS
20	30.12	18.88	0.782	12.91	0.390
40	33.33	16.02	0.826	10.72	0.305
60	36.77	14.01	0.875	9.21	0.220
80	38.41	12.88	0.902	8.40	0.160
100	39.03	11.34	0.921	7.85	0.118

Table 3. Comparative evaluation of total registration error (TRE) for different motion compensation (MC) algorithms

Motion compensation method	Mean TRE [voxels]	Std. Dev.	P-value
No motion compensation	4.91	2.65	<0.002
Wachinger et al.	2.44	0.61	<0.002
Yang et al.	2.70	0.30	<0.002
MIND	1.91	0.12	<0.002
SyN (ANTs)	1.78	0.15	<0.002
VoxelMorph	1.69	0.13	<0.002
DLIR	1.61	0.14	<0.002
DDFReg	1.56	0.11	<0.002
Proposed method	1.44	0.10	<0.002

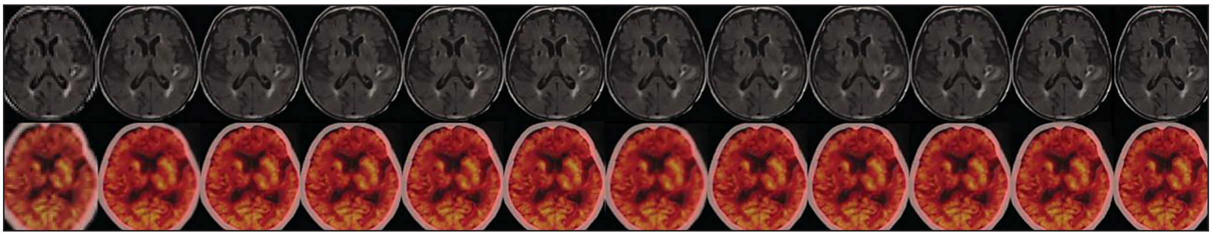

Figure 4. Whole-slice comparison of CT-PET reconstructions using 12 super-resolution techniques. Top row presents full-resolution CT-PET overlays; bottom row provides zoomed-in views to highlight tissue contrast and lesion localization. The proposed RED-WGAN yields superior structural clarity and functional consistency across modalities

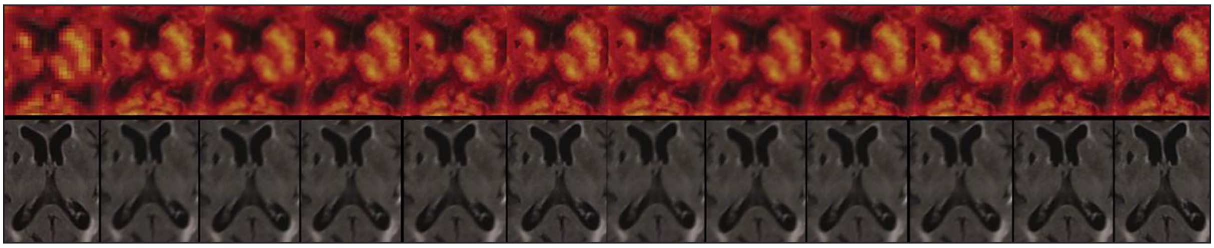

Figure 5. Zoomed region comparison of CT-PET reconstructions. Top row shows high-magnification PET images; bottom row displays the corresponding structural overlays. The proposed RED-WGAN demonstrates enhanced lesion detectability, improved noise suppression, and superior anatomical alignment compared to competing methods

was quantitatively validated using total registration error (TRE), reported in Table 3. RED-WGAN achieved the lowest mean TRE (1.44 voxels) and standard deviation (0. 10), outperforming classical and recent deep learning-based registration models.

Visual comparisons from additional scans – including brain and thoracic PET/CT data – are presented in Figures 4 and 5. These results further highlight RED-WGAN's ability to recover high-resolution detail from sparsely sampled or motion-degraded acquisitions.

The statistical significance of improvements was verified using paired t-tests across all sampling conditions and methods ($p<0.01$). Bowker symmetry testing further confirmed the stability and consistency of key metrics across repeated simulation trials.

In summary, RED-WGAN not only improves image fidelity and perceptual realism, but does so under challenging clinical constraints such as compressed sensing and respiratory motion. Its integration of deformable motion correction and GAN-based reconstruction enables accurate recovery of anatomical structures under low-dose and undersampled regimes, offering practical clinical advantages in terms of acquisition speed, patient safety, and diagnostic confidence.

CONCLUSIONS

This study has presented RED-WGAN, a novel and technically sophisticated framework for CT-PET image reconstruction, which effectively bridges the gap between acquisition limitations and diagnostic quality. The method was designed to address a convergence of critical challenges in hybrid imaging—namely, resolution disparity between CT and PET modalities, motion artifacts, data sparsity, and the increasing need for low-dose imaging protocols. By unifying deformable motion correction, residual deblurring, and Transformer-based denoising within a generative adversarial network, RED-WGAN achieves a level of performance that sets a new benchmark for the field.

Extensive quantitative and qualitative evaluations have confirmed the superiority of the proposed framework across multiple experimental configurations. When benchmarked against twelve leading super-resolution and reconstruction methods, RED-WGAN consistently achieved higher fidelity in terms of PSNR, SSIM, MAE, and LPIPS, while also delivering the lowest registration error (TRE), demonstrating a significant improvement in anatomical alignment and perceptual image quality. Importantly, these results were maintained even under extreme data undersampling (as low as 40%), indicating the robustness and adaptability of the model under real-world imaging constraints.

The integration of Ridgelet-domain sparsity and compressed sensing principles into the learning pipeline further enables RED-WGAN to exploit global and local correlations in both anatomical and functional data. This is particularly valuable in clinical scenarios where radiation dose, scan time, or patient cooperation must be minimized—such as pediatric imaging, emergency diagnostics, and follow-up oncology studies. Moreover, the use of Transformer-based modules enhances the model's ability to capture long-range dependencies and global contextual features, which are essential for recovering fine anatomical structures and resolving texture-level ambiguities in degraded inputs.

The deformable registration component embedded within RED-WGAN represents another major innovation, enabling effective compensation for respiratory and patient-induced motion without relying on external gating or hardware. This contributes to improved inter-modality congruence between CT and PET volumes, enhancing the interpretability of fused images and facilitating more accurate lesion localization and staging.

Beyond technical performance, the model has been validated on both synthetic phantom datasets and diverse in vivo clinical CT-PET volumes, ensuring generalizability across anatomical sites and imaging conditions. The architecture is scalable to 3D volumetric data and is compatible with existing clinical infrastructure, allowing for future integration into cloud-based or edge-deployed diagnostic platforms.

REFERENCES

1. Bruna J., Sprechmann P., LeCun Y. Super-resolution with deep convolutional sufficient statistics. Proc. ICLR 2015. https://doi.org/10.48550/arXiv.1511.05666
2. Isola P., Zhu J.Y., Zhou T., Efros A.A. Image-to-image translation with conditional adversarial networks. Proc. IEEE Conf. CVPR, Honolulu, HI, USA, 2017; 5967–5976. https://doi.org/10.1109/CVPR.2017.632
3. Ledig C. et al. Photo-realistic single image super-resolution using a generative adversarial network. Proc. IEEE Conf. CVPR, Honolulu, HI, USA, 2017; 4681–4690. https://doi.org/10.1109/CVPR.2017.19
4. Malczewski K. Super-resolution with compressively sensed MR/PET signals at its input. Informatics in Medicine Unlocked 2020; 20: 100312. https://doi.org/10.1016/j.imu.2020.100312
5. Jin K.H., Ye J.C. Annihilating filter-based low-rank Hankel matrix approach for image inpainting. IEEE Trans. Image Process. 2015; 24(12): 3498–3511. https://doi.org/10.1109/TIP.2015.2484298
6. Jin K.H. et al. Generalized Hankel matrix decomposition for low-rank signal recovery. IEEE Trans. Image Process. 2016; 25(6): 2739–2750. https://doi.org/10.1109/TIP.2016.2547738
7. Dong C., Loy C.C., He K., Tang X. Image super-resolution using deep convolutional networks. IEEE Trans. Pattern Anal. Mach. Intell. 2016; 38(2): 295–307. https://doi.org/10.1109/TPAMI.2015.2439281
8. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014. https://doi.org/10.48550/arXiv.1409.1556
9. Johnson J., Alahi A., Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. Proc. ECCV 2016; 694–711. https://doi.org/10.1007/978-3-319-46475-6_43
10. Donoho D.L. Compressed sensing. IEEE Trans. Inf. Theory 2006; 52(4): 1289–1306. https://doi.org/10.1109/TIT.2006.871582
11. Lustig M., Donoho D., Pauly J.M. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn. Reson. Med. 2007; 58(6): 1182–1195. https://doi.org/10.1002/mrm.21391
12. Matej S., Lewitt R.M. Efficient 3D iterative reconstruction using a projection-space representation. IEEE Trans. Med. Imaging 2004; 23(5): 553–564. https://doi.org/10.1109/TMI.2004.827684
13. Brady M., Muftuoglu O. Fast Discrete Radon Transform: Theory and Implementation. Int. J. Imaging Syst. Technol. 1998; 9(1): 43–52. https://doi.org/10.1002/(SICI)1098-1094(1998)9:1<43::AID-IMA2>3.0.CO;2-N
14. Surti S., Karp J.S. Design considerations for a limited-angle, dedicated breast TOF PET scanner. Phys. Med. Biol. 2009; 54(14): 4141–4157. https://doi.org/10.1088/0031-9155/54/14/011
15. Knoll F. et al. Joint reconstruction of simultaneously acquired CT-PET data with multi-sensor compressed sensing based on a mutual sparsity

constraint. EJNMMI Phys. 2014; 1: 1. https://doi.org/10.1186/2197-7364-1-1

16. Antoch G., Bockisch A. Combined PET/MRI: A new dimension in whole-body oncology imaging. EJNMMI 2008; Suppl 1: S113–S120. https://doi.org/10.1007/s00259-008-0819-1
17. Goodfellow I. et al. Generative adversarial nets. Adv. Neural Inf. Process. Syst. (NeurIPS) 2014; 27: 2672–2680. https://doi.org/10.5555/2969033.2969125
18. Hu X. et al. Meta-SR: A magnification-arbitrary network for super-resolution. Proc. CVPR 2019; 1575–1584. https://doi.org/10.1109/CVPR.2019.00162
19. McGibney G. et al. Quantitative evaluation of several partial Fourier reconstruction algorithms used in MRI. Magn. Reson. Med. 1993; 30(5): 51–59. https://doi.org/10.1002/mrm.1910300508
20. Cheryauka A.B. et al. MRI diffusion tensor reconstruction with PROPELLER data acquisition. Magn. Reson. Imaging 2004; 22(2): 139–148. https://doi.org/10.1016/j.mri.2003.10.002
21. Lim B. et al. Enhanced deep residual networks for single image super-resolution. Proc. CVPR Workshops 2017; 1132–1140. https://doi.org/10.1109/CVPRW.2017.151
22. Zhang Y. et al. Image super-resolution using very deep residual channel attention networks. Proc. ECCV 2018; 294–310. https://doi.org/10.1007/978-3-030-01234-2_18