



ARTICLE

# Hybrid Laplacian-DoG: Noise-Preserving 3D FDG-PET Contrast Enhancement for Improved MCI Detection

Ovidijus Grigas\*  and Rytis Maskeliūnas 

Department of Software Engineering, Kaunas University of Technology, Kaunas, Lithuania

\*Corresponding Author: Ovidijus Grigas. Email: o.grigas@ktu.edu

Received: 07 December 2025; Accepted: 23 March 2026; Published: 27 April 2026

**ABSTRACT:** Early detection of Mild Cognitive Impairment (MCI) with FDG-PET is essential for timely Alzheimer's disease intervention. However, PET image quality is limited by low spatial resolution, partial volume effects, and Poisson noise. Standard enhancement methods, such as Bilateral filtering or Contrast Limited Adaptive Histogram Equalization (CLAHE), can increase contrast but often introduce heavy noise or distort image texture, while deep learning methods may produce hallucinated structures. We propose a fully data-adaptive, non-learned 3D enhancement framework whose output is deterministic for a given input volume, that combines Laplacian-based local contrast modulation with a gradient-gated Difference-of-Gaussians (DoG) detail injector. This hybrid design sharpens anatomical boundaries while keeping noise amplification near unity in uniform regions. The method enhances structure only where true radiotracer gradients are present. We evaluated the approach on a large Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort ( $N = 1928$ ). Quantitative results show that the method increases contrast without adding noise, achieving a Noise Gain of 1.01 (vs. 1.28 for Bilateral filtering) and a high Edge Preservation Index (0.981). In downstream classification experiments across multiple deep learning architectures, the greatest improvement was observed in MobileNetV4 on the axial plane, where mean accuracy increased from 93% to 96%. Overall, the proposed gradient-gated hybrid enhancement provides a reliable PET pre-processing strategy. By recovering subtle metabolic patterns without amplifying noise, it strengthens the sensitivity of automated MCI diagnostic systems.

**KEYWORDS:** Positron emission tomography; image enhancement; mild cognitive impairment; classification

## 1 Introduction

Alzheimer's Disease (AD) remains a major challenge for global public health, and Mild Cognitive Impairment (MCI) constitutes a transitional stage in which therapeutic interventions are most likely to alter disease progression [1]. Although magnetic resonance imaging (MRI) captures structural changes, it often does so only after significant neuronal loss has occurred [2]. In contrast,  $^{18}\text{F}$ -fluorodeoxyglucose Positron Emission Tomography ( $^{18}\text{F}$ -FDG-PET) provides a functional map of cerebral glucose metabolism, capable of identifying hypometabolic biomarkers before structural changes become visible [3]. However, the diagnostic value of FDG-PET is limited by the physical properties of the modality itself. PET images have a relatively low spatial resolution (about 4–6 mm), a low Signal-to-Noise Ratio (SNR), and are affected by Partial Volume Effects (PVE) [4]. As a result, signals from small cortical structures tend to mix with those of nearby tissue, which blurs important boundaries. For MCI, where metabolic changes are mild, blurring can hide early abnormalities and reduce diagnostic accuracy [5].

Enhancing PET images is challenging because increasing contrast often increases noise [6]. Traditional methods attempt to address this trade-off but still have notable limitations. In routine clinical workflows,

Gaussian smoothing is applied to reduce Poisson noise, but it also increases partial-volume effects and blurs anatomical boundaries [7]. Iterative approaches such as Anisotropic Diffusion aim to smooth noise within regions while preserving edges, but can produce unrealistic “plastic” textures or staircase artifacts that negatively affect radiomic measures [8]. Wavelet-based denoising methods separate noise and signal across frequency bands, though they often introduce ringing around high-contrast metabolic regions [9]. Histogram-based methods like Contrast Limited Adaptive Histogram Equalization (CLAHE) can improve local contrast, but their intensity redistribution can distort quantitative uptake values (Standardized Uptake Values—SUV), introduce block-like artifacts, and saturate high-uptake areas, reducing the reliability of the scan [10].

In recent years, Deep Learning (DL) approaches such as Super-Resolution GANs (SRGANs) [11], CycleGANs [12], and U-Net-based models [13] have been widely adopted for PET enhancement. These methods can generate detailed textures and reduce noise, often outperforming classical techniques in perceptual quality metrics. However, several issues limit their clinical applicability. The most serious concern is “hallucination”, where GAN-based models produce realistic-looking but artificial anatomical structures that are not present in the patient’s scan [14]. In addition, these models require large paired datasets (e.g., low-dose/high-dose PET pairs), which are rarely available in real clinical settings [15]. Their performance also tends to drop when applied to data from different scanners, as small variations in reconstruction protocols or hardware can disrupt generalization [16]. Finally, the “black box” nature of deep neural networks raises concerns about interpretability and reliability. Clinicians often hesitate to trust the pixel-level modifications produced by models whose decision processes cannot be easily explained [17].

To address these limitations without relying on “black-box” or unstable deep learning models, we propose a data-driven 3D enhancement framework. The method combines a Laplacian operator for edge detection with a Difference-of-Gaussians (DoG) filter for band-pass detail extraction. A gradient-gating mechanism ensures that sharpening occurs only along true anatomical boundaries, while uniform regions and their native noise characteristics remain unchanged. This design produces a parameter-efficient algorithm that improves local metabolic contrast while keeping noise amplification close to unity ( $NG \approx 1.0$ ).

The main contributions of this work are summarized as follows:

1. **Hybrid 3D Enhancement Framework:** We present a hybrid algorithm that merges Laplacian-based contrast modulation with gradient-gated DoG detail injection. This formulation effectively separates contrast enhancement from noise amplification, overcoming a key limitation of classical filtering.
2. **Safety-Oriented Evaluation Metrics:** We propose and apply radiomic safety metrics Noise Gain (NG) and Edge Preservation Index (EPI) to demonstrate that the proposed method avoids texture corruption and artifacts commonly introduced by Bilateral filtering.
3. **Clinical Validation on MCI Detection:** Using a large ADNI cohort ( $N = 1928$ ), we show that this lightweight pre-processing step improves downstream MCI classification by more than 3% across modern deep learning architectures (including MobileNetV4 and MambaOut), reaching a mean accuracy of 96%. These results highlight that careful signal processing can significantly enhance deep learning performance.

## 2 Related Works

Enhancement of FDG-PET image quality has been approached through both classical signal-processing pipelines and modern deep learning techniques. In the following, we summarize recent open-access studies that address the balance between resolution, noise reduction, and quantitative reliability in PET imaging.

Flaus et al. [18] introduced a deep learning framework to improve the visibility of focal epilepsy lesions using high-quality simulated PET phantoms as ground truth. A ResNet model was trained to map standard

low-quality FDG-PET scans to sharpened, high-resolution output. The method improved quantitative metrics (Peak Signal-to-Noise Ratio–PSNR, Structural Similarity Index Measure–SSIM) and substantially increased the detection rate of subtle cortical hypometabolism (38% to 75%), with greater reader confidence.

Chaudhari et al. [19] developed a Convolutional Neural Network (CNN)-based low-count PET enhancement method for whole-body oncology imaging. Using a U-Net-style architecture, the model restored image quality in scans acquired at one-quarter of the standard dose. Blinded readers across multiple centers rated the enhanced scans as not inferior to full-dose reconstructions. Importantly, SUV measurements, lesion detectability, and diagnostic performance were preserved (sensitivity 0.94, specificity 0.98).

Song et al. [20] applied a Generative Adversarial Network for PET super-resolution. By incorporating adversarial loss, the GAN produced sharper cortical boundaries and more realistic uptake textures than conventional CNNs. Although perceptual improvements were notable, the authors emphasized the need for careful validation to avoid introducing hallucinated features.

Hashimoto et al. [21] explored the use of the deep image prior (DIP) for PET denoising. Instead of relying on external datasets, a convolutional generator was optimized directly on each subject's dynamic PET series, using the high-count frame of each scan as implicit supervision. The method preserved temporal uptake patterns and outperformed Gaussian and guided filtering. A later 3D U-Net variant further improved denoising quality while avoiding excessive smoothing.

Jiang et al. [22] proposed TriPLET, an end-to-end multi-domain framework that processes PET data in the projection, frequency, and image domains. The pipeline couples a Transformer-based sinogram denoiser, a wavelet U-Net reconstructor, and a GAN discriminator to enforce multi-level consistency. TriPLET achieved state-of-the-art results on real low-dose PET data, recovering standard-dose image quality from quarter-dose inputs with improved SNR and structural fidelity.

Xue et al. [23] introduced LCPR-Net, a hybrid reconstruction–super-resolution model based on domain-transform CNNs and CycleGAN training. Operating directly on low-count sinograms, the network reconstructs full-count PET images with a cyclic consistency constraint that reduces the risk of hallucination. LCPR-Net achieved the highest PSNR/SSIM and the lowest error across baselines, while outperforming conventional iterative methods in both speed and image quality.

Yoshimura et al. [24] proposed a Residual Dense Network for super-resolution of half-duration FDG-PET scans. Trained on paired low-count and full-count data from 108 subjects, the model produced images with markedly improved contrast and clarity. Visual assessments confirmed that super-resolved PET images closely matched the quality of standard full-dose scans.

Chen et al. [25] presented Deep Progressive Learning (DPL), an AI-driven reconstruction algorithm integrated directly into the PET reconstruction pipeline. DPL improved image quality across all body mass index (BMI) groups compared with standard Ordered Subset Expectation Maximization (OSEM), resulting in better lesion visibility and increased diagnostic confidence. The method demonstrated consistent gains across patients' body compositions.

To provide a structured overview of the research area, [Table 1](#) summarizes the key characteristics of classical signal-processing methods and modern deep learning approaches for PET image enhancement. This comparison highlights the trade-offs that motivate the design of the proposed framework.

As shown in [Table 1](#), classical methods offer transparency and quantitative stability but struggle to enhance fine structural detail without amplifying noise. Deep learning methods can achieve superior perceptual quality but introduce risks of hallucination, dataset dependency, and limited generalization. The proposed Hybrid Laplacian-DoG framework combines the interpretability and noise stability of classical approaches with targeted structural enhancement, without requiring training data or introducing learned artifacts.

**Table 1:** Systematic comparison of classical and deep learning PET enhancement approaches.

Characteristic	Classical methods	Deep learning methods
Workflow	Post-reconstruction filtering	End-to-end or reconstruction-integrated
Training Data Required	None	Large paired/unpaired datasets
Hallucination Risk	None	Moderate to high (especially GANs)
Quantitative Fidelity (SUV)	Generally preserved	May be altered by learned mappings
Cross-Scanner Generalization	High (parameter-driven)	Limited without retraining/fine-tuning
Interpretability	Fully transparent	Low (“black box”)
Noise Handling	Often amplifies or over-smooths	Can suppress but may hallucinate texture
Computational Cost	Low (seconds per volume)	High (training: hours/days; inference: variable)
Clinical Deployability	Immediate, no infrastructure needed	Requires GPU, validated models, pipelines

### 3 Materials and Methods

#### 3.1 Dataset and Preprocessing

The experimental evaluation was performed using ADNI FDG-PET data (available online: [adni.loni.usc.edu](http://adni.loni.usc.edu)). The aim of ADNI is to determine whether MRI, PET, biological markers, and cognitive assessments can jointly characterize the progression of MCI and early Alzheimer’s disease. FDG-PET scans were obtained from ADNI-1, ADNI-GO, and ADNI-2 phases, using ADNI’s baseline diagnostic labels.

In this study, we focused on FDG-PET, as cerebral glucose metabolism is a sensitive marker of neuronal dysfunction associated with cognitive decline. To support a balanced and unbiased assessment of the proposed enhancement method in downstream classification experiments, we constructed a dataset consisting of 964 Cognitively Normal (CN) and 964 MCI scans, yielding a total of  $N = 1928$  3D volumes.

Subjects were selected according to their baseline diagnostic status, as defined by standard ADNI criteria. CN subjects showed no memory complaints, demonstrated normal memory performance adjusted for age and education, had a CDR score of 0, and showed no evidence of significant neurological or psychiatric disease. Subjects classified as MCI met ADNI criteria, including subjective memory concern, objective memory impairment on standardized neuropsychological tests, largely preserved activities of daily living, and absence of dementia. All selections were performed at the subject level to avoid data leakage.

[Table 2](#) summarizes the demographic and clinical characteristics of the study cohort.

All scans underwent a standardized preprocessing pipeline consisting of the following sequential steps:

1. **Brain extraction:** Non-brain tissue was removed using SynthStrip [26], a learning-based skull-stripping tool that generalizes across modalities without requiring modality-specific retraining.
2. **Spatial normalization:** Each brain-extracted volume was registered to the Montreal Neurological Institute (MNI) 152 FDG-PET template using FSL FLIRT [27]. The resulting volumes were resampled to an isotropic voxel resolution of  $2 \times 2 \times 2 \text{ mm}^3$ .
3. **Intensity masking:** A binary brain mask  $M$  was generated by retaining voxels with strictly positive values exceeding a noise threshold defined as  $0.002 \times P_{99}$ , where  $P_{99}$  is the 99th percentile of valid intensities in the scan. This step removes residual non-brain background and low-intensity noise voxels.

No additional smoothing, intensity normalization (e.g., global mean scaling), or partial volume correction was applied prior to enhancement, ensuring that the evaluation reflects the method’s ability to operate on minimally processed clinical data.

**Table 2:** Demographic and clinical characteristics of the study cohort. MMSE = Mini-Mental State Examination; CDR = Clinical Dementia Rating; CDR-SB = CDR Sum of Boxes; FAQ = Functional Activities Questionnaire; F = female; M = male.

Characteristic	CN	MCI
Scans selected	964	964
Age (years)	75.1 ± 6.3	73.8 ± 7.6
Sex (F/M)	50.4%/49.6%	39.5%/60.5%
Education (years)	16.1 ± 2.8	15.8 ± 2.9
MMSE	29.0 ± 1.2	27.5 ± 1.8
CDR Global	0.0 ± 0.0	0.5 ± 0.06
CDR-SB	0.03 ± 0.13	1.53 ± 0.94
FAQ Total	0.2 ± 0.6	3.3 ± 4.1
Volume dimensions (voxels)	91 × 109 × 91	
Voxel size (mm <sup>3</sup> )	2.0 × 2.0 × 2.0	

### 3.1.1 Computational Environment

All experiments were conducted on a Linux-based workstation running Ubuntu 22.04, a single NVIDIA RTX 4090 GPU, paired with an AMD Ryzen 9 5900X CPU and 32 GB of system memory. The processing and evaluation pipeline was implemented in Python (v3.10), using NumPy and SciPy for numerical computation, NiBabel for neuroimaging data handling, and scikit-image for image processing operations. Brain extraction was performed using SynthStrip [26], and spatial normalization was carried out with FSL [27].

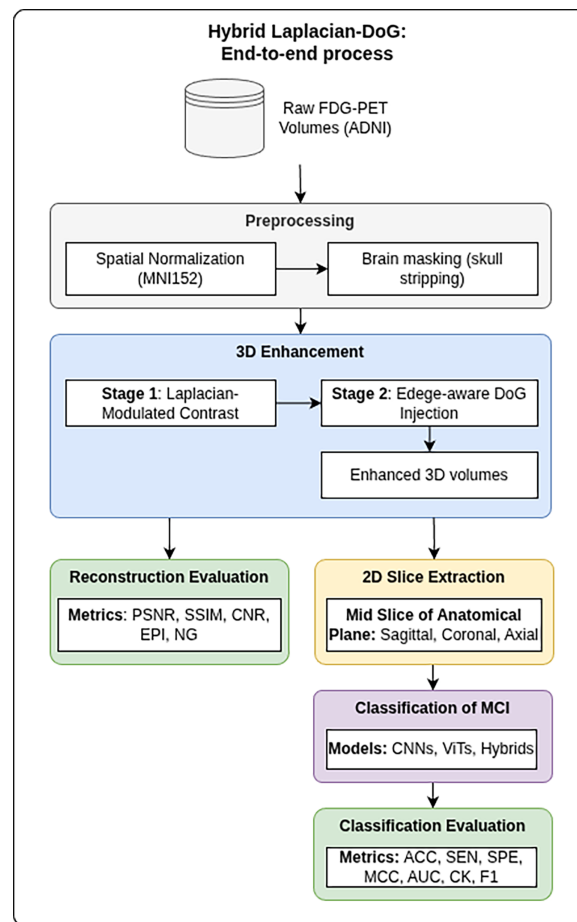
The proposed enhancement framework is fully deterministic and implemented using standard convolution and point-wise operations. Processing a single 3D FDG-PET volume requires approximately 0.6–1.0 s on CPU, depending on I/O overhead. Over 1000 repeated runs on a representative FDG-PET volume yielded an average runtime of  $0.84 \pm 0.16$  s. No training or model inference is required. This runtime is substantially lower than that of deep learning based enhancement or reconstruction methods and enables practical deployment in high-throughput clinical workflows.

### 3.2 Proposed Hybrid Enhancement Framework

We propose a hybrid 3D enhancement framework that combines Laplacian-modulated contrast amplification with an edge-aware DoG detail injector. The approach operates sequentially, first enhancing the broad structural contrast, and then selectively adding high-frequency detail. Let  $G_\sigma(\mathbf{x})$  denote a 3D Gaussian kernel with standard deviation  $\sigma$ , and let  $*$  represent convolution. A small constant  $\varepsilon = 10^{-6}$  is included in all division operations to maintain numerical stability. This value was chosen to be several orders of magnitude below the minimum meaningful intensity in normalized FDG-PET volumes (which are typically on the order of  $10^{-2}$  to  $10^0$  after percentile-based normalization), ensuring that  $\varepsilon$  prevents division-by-zero errors without influencing the computed ratios in any physiologically relevant regime. The specific choice of  $10^{-6}$  follows standard numerical practice in image processing and is consistent with single-precision floating-point resolution ( $\sim 10^{-7}$ ).

The proposed enhancement framework is intended to be used as a pre-processing step for automated image analysis pipelines rather than as a standalone reconstruction or diagnostic tool. Its design prioritizes structural fidelity and noise preservation to improve the reliability of downstream computational tasks, such as classification or feature extraction.

Fig. 1 illustrates the conceptual enhancement algorithm. For clarity, we summarize here the complete experimental workflow used in this study. Each FDG-PET volume undergoes: spatial normalization and brain masking; enhancement using the proposed hybrid method or comparison baselines; extraction of 2D mid-slices in sagittal, coronal, and axial planes; input preparation and augmentation for classification networks; and quantitative evaluation using reconstruction and classification metrics. Enhancement is applied independently to each 3D volume prior to any learning-based processing, and no information from the classification stage is used to tune enhancement parameters.



**Figure 1:** End-to-end processing pipeline used in this study. Raw FDG-PET volumes are spatially normalized and brain-masked, enhanced using the proposed hybrid or baseline methods, and then converted into 2D mid-slices for classification. Reconstruction quality metrics are computed directly on enhanced 3D volumes, while classification performance is evaluated on the extracted slices.

For clarity, all weighting terms are defined voxel-wise and indexed by spatial location  $\mathbf{x}$ . Some weights depend on local intensity values rather than spatial derivatives. In such cases, the notation  $W(\mathbf{x})$  denotes a spatially indexed function whose value is computed from the intensity at that voxel, e.g.,  $W_{int}(\mathbf{x}) \equiv W_{int}(V(\mathbf{x}))$  or  $W_{int}(V_{lap}(\mathbf{x}))$  where specified. This notation emphasizes voxel-wise application rather than an independent spatial variable.

### 3.2.1 Stage 1: Laplacian-Modulated Contrast

Conventional contrast enhancement techniques, such as histogram equalization, often perform poorly in PET volumes because they indiscriminately amplify noise in low-uptake regions and oversaturate high-intensity metabolic hotspots [28]. In contrast, the Laplacian module functions as a selective amplifier: local gain  $g(\mathbf{x})$  is directly related to spatial edge strength, ensuring that contrast increases only where there is a meaningful anatomical structure. The accompanying intensity weighting restricts the enhancement to the metabolic range characteristic of gray matter, thereby avoiding amplification of background noise and minimizing ringing artifacts around regions of high uptake. The procedure is formally defined below.

The first stage decomposes the input volume  $V$  into a base layer  $V_{base} = V * G_{\sigma_{base}}$  and a detail layer  $D = V - V_{base}$ , with  $\sigma_{base} = 1.0$ . We amplify the detail layer using a spatially varying gain map  $g(\mathbf{x})$ .

The gain map relies on two weighting components: edge strength and signal intensity. First, we compute the Laplacian of Gaussian (LoG) response to detect structural boundaries:

$$L(\mathbf{x}) = |\nabla^2(G_{\sigma_{edge}} * V)(\mathbf{x})| \quad (1)$$

where  $\sigma_{edge} = 3.0$ , corresponding to approximately 6mm Full Width at Half Maximum (FWHM), which matches the typical spatial resolution of PET scanners and the scale of cortical gray–white matter transitions. To define the edge weight  $W_{edge}$ , we normalize  $L(\mathbf{x})$  by its 90th percentile ( $P_{90}$ ):

$$W_{edge}(\mathbf{x}) = \left[ \min \left( 1, \frac{L(\mathbf{x})}{P_{90}(L) + \epsilon} \right) \right]^\gamma \quad (2)$$

where  $\gamma = 0.8$  applies mild compression that boosts mid-strength edges (such as subtle cortical boundaries) relative to dominant high-contrast edges, preventing the enhancement from being dominated by a few strong structures. We enforce a strict range  $[0, 1]$ .

Second, to avoid amplifying noise in low-uptake regions and to prevent saturation in high-intensity metabolic areas, we introduce an intensity weighting term  $W_{int}$ . Voxel intensities are first normalized to the  $[0, 1]$  range using a metabolic window defined by the 20th ( $I_{lo}$ ) and 85th ( $I_{hi}$ ) percentiles of valid brain intensities. The 20th percentile approximately separates low-uptake white matter and CSF from metabolically active tissue, while the 85th percentile captures the upper range of cortical gray matter uptake without being distorted by occasional high-intensity voxels:

$$I_{norm}(\mathbf{x}) = \text{clip} \left( \frac{V(\mathbf{x}) - I_{lo}}{I_{hi} - I_{lo} + \epsilon}, 0, 1 \right) \quad (3)$$

$W_{int}$  is then modeled as a Gaussian curve centered at 0.5 (mid-range intensity, corresponding to the gray matter metabolic range) with a spread of  $\sigma_w = 0.25$ , which covers approximately  $\pm 1$  standard deviation of the normalized intensity range and ensures that both low-uptake background ( $I_{norm} \ll 0.5$ ) and high-uptake hotspots ( $I_{norm} \gg 0.5$ ) are progressively downweighted:

$$W_{int}(\mathbf{x}) = \exp\left(-\frac{(I_{norm}(\mathbf{x}) - 0.5)^2}{2\sigma_w^2}\right) \quad (4)$$

The final Laplacian-enhanced volume  $V_{lap}$  is computed as

$$V_{lap}(\mathbf{x}) = V_{base}(\mathbf{x}) + (1 + k \cdot W_{edge}(\mathbf{x}) \cdot W_{int}(\mathbf{x})) \cdot (V(\mathbf{x}) - V_{base}(\mathbf{x})) \quad (5)$$

where  $k = 2.0$  determines the maximum contrast gain. At full weighting ( $W_{edge} = W_{int} = 1$ ), the detail layer is amplified by a factor of  $1 + k = 3$ , which provides substantial contrast improvement without introducing ringing or oversaturation at typical PET intensity levels.

### 3.2.2 Stage 2: Edge-Aware DoG Injection

Although Stage 1 improves local contrast, PET volumes still lack fine structural definition due to their inherently low spatial resolution. Stage 2 compensates for this by injecting high-frequency detail extracted using a DoG filter. A key component of this hybrid design is the gradient-gate term  $W_{grad}$ . Conventional sharpening methods amplify noise and structural edges indiscriminately, making it difficult to distinguish tissue boundaries from background grain. In contrast, the gradient gate suppresses sharpening in locally homogeneous regions, ensuring that detail enhancement is applied only where genuine anatomical gradients exist. It enables a strong delineation of cortical folds and metabolic boundaries without degrading the signal-to-noise ratio elsewhere.

The second stage introduces targeted band-pass detail in  $V_{lap}$ . The DoG response is computed using a fine scale  $\sigma_s = 0.7$  and a coarse scale  $\sigma_l = 3.0$ :

$$D_{DoG}(\mathbf{x}) = (V_{lap} * G_{\sigma_s})(\mathbf{x}) - (V_{lap} * G_{\sigma_l})(\mathbf{x}) \quad (6)$$

Adding directly  $D_{DoG}$  would ordinarily amplify high-frequency noise. To avoid this, the DoG contribution is modulated by the local gradient magnitude. Importantly, the gradient is computed on a fine-scale smoothed version of the volume  $V_s = V_{lap} * G_{\sigma_s}$ , which stabilizes the estimate and reduces sensitivity to voxel-wise noise:

$$|\nabla V_s(\mathbf{x})| = \sqrt{\left(\frac{\partial V_s}{\partial x}\right)^2 + \left(\frac{\partial V_s}{\partial y}\right)^2 + \left(\frac{\partial V_s}{\partial z}\right)^2} \quad (7)$$

The gradient suppression mask  $W_{grad}$  is defined using an inverted Gaussian function, with the parameter  $\tau = 0.01$  controlling the sensitivity to locally flat regions:

$$W_{grad}(\mathbf{x}) = 1 - \exp\left(-\left(\frac{|\nabla V_s(\mathbf{x})|}{\max(\tau, \epsilon)}\right)^2\right) \quad (8)$$

The gradient-gating function is designed to act as a soft anatomical edge detector rather than as an explicit model of PET noise statistics. In FDG-PET, Poisson noise is approximately signal-dependent and largely uncorrelated at the voxel level, whereas true anatomical boundaries manifest themselves as coherent gradients that persist after light smoothing. The exponential form of  $W_{grad}$  provides a smooth and monotonic transition between flat regions—where high-frequency detail should not be injected—and regions with sufficiently strong gradients to indicate plausible anatomical structure. The parameter  $\tau$  serves

as a scale threshold that separates noise-dominated gradients from structure-dominated gradients after fine-scale smoothing. This formulation avoids hard thresholding, maintains differentiability, and ensures that enhancement is suppressed in regions where the gradient magnitude matches that of pure Poisson noise.

The final hybrid volume  $V_{hybrid}$  is produced by re-inserting the weighted DoG detail, that is modulated by both the gradient mask and the intensity-based weighting computed on  $V_{lap}$ , using a mixing factor of  $\alpha = 0.6$ :

$$V_{hybrid}(\mathbf{x}) = V_{lap}(\mathbf{x}) + \alpha \cdot (W_{grad}(\mathbf{x}) \cdot W_{int}(V_{lap}(\mathbf{x}))) \cdot D_{DoG}(\mathbf{x}) \quad (9)$$

Finally,  $V_{hybrid}$  is linearly rescaled so that its global median matches the global median of the original input  $V$ .

### 3.2.3 Parameter Selection

The proposed framework includes a small set of scale and weighting parameters chosen to align with the spatial resolution and intensity characteristics of FDG-PET rather than to optimize any specific downstream metric. The Gaussian scales were selected to reflect physiologically plausible structural ranges:  $\sigma_{base} = 1.0$  applies light smoothing at approximately one voxel and preserves gyral anatomy;  $\sigma_{edge} = 3.0$  and  $\sigma_l = 3.0$  correspond to spatial extents on the order of cortical thickness, capturing broad gray-white matter transitions; and the fine-scale  $\sigma_s = 0.7$  is used solely to stabilize gradient and DoG estimates at sub-voxel resolution. The intensity-weighting width  $\sigma_w = 0.25$  defines a window around mid-range gray-matter uptake in the normalized  $[0, 1]$  space, ensuring that low-uptake background and high-uptake hotspots are downweighted during contrast modulation.

The modulation parameters  $k = 2.0$ ,  $\alpha = 0.6$ ,  $\gamma = 0.8$ , and  $\tau = 0.01$  influence the shape of the weighting functions rather than their spatial location. In particular,  $k$  constrains the maximum Laplacian-based contrast gain;  $\alpha$  limits the contribution of DoG detail injection;  $\gamma$  applies a soft emphasis to mid-strength edges without amplifying noise; and  $\tau$  defines the transition threshold between flat and non-flat regions in the gradient gate.

Percentile-based quantities such as  $P_{90}$  for Laplacian normalization,  $P_{20}$  and  $P_{85}$  for defining the metabolic window, and  $P_{30}$  for identifying flat regions are computed adaptively for each volume. This ensures invariance to inter-scanner intensity scaling, which is essential in FDG-PET where absolute intensity values are not standardized across acquisitions. While fixed thresholds would be highly sensitive to scanner calibration and dose, percentiles provide a robust, distribution-based normalization.

These percentile anchors are deterministic functions of the input volume: given a specific scan, the same percentiles and enhancement are always obtained. As such, they function not as tunable hyperparameters but as stable normalization references tied directly to each subject's empirical intensity distribution. Because their role is normalization rather than optimization, they are not subject to sensitivity tuning in the same sense as modulation parameters (e.g.,  $k$ ,  $\alpha$ ,  $\tau$ ). We therefore restrict sensitivity analysis to the primary modulation parameters and present the results in [Section 4.1.2](#).

For all comparison methods, parameters were fixed across the entire dataset and were not tuned on a per-image, per-subject, or task-specific basis. Parameter values were selected based on commonly used settings reported in the literature and preliminary visual inspection to ensure stable behavior on FDG-PET volumes, rather than to optimize any quantitative metric. No dataset-level optimization or label-driven tuning was performed for any baseline method.

Specifically, unsharp masking was implemented using a Gaussian blur with  $\sigma = 1.5$  and scaling factor  $\lambda = 0.5$ ; slice-wise bilateral filtering used spatial and radiometric parameters  $\sigma_s = 2.0$  and  $\sigma_r = 0.1$ ; CLAHE

was applied slice-wise using  $8 \times 8$  tiles. These settings reflect standard practice and were kept identical for all scans to ensure a fair and reproducible comparison.

### 3.3 Comparative Analysis

We compare the proposed method with three established enhancement techniques. In addition, because each component of the framework can be applied independently, evaluating the Laplacian-only, DoG-only, and full Hybrid variants provides an explicit ablation study of the method's constituent stages. This allows us to isolate the contribution of the structural Laplacian backbone, the DoG-based detail injection, and their combined effect.

#### 3.3.1 3D Unsharp Masking

Traditional linear unsharp masking enhances the high-frequency content by subtracting a blurred version of the volume from the original. For an input volume  $V$  and a Gaussian kernel  $G_\sigma$  with  $\sigma = 1.5$ , the enhanced output  $V_{unsharp}$  is computed as:

$$V_{unsharp}(\mathbf{x}) = V(\mathbf{x}) + \lambda (V(\mathbf{x}) - [G_\sigma * V](\mathbf{x})) \quad (10)$$

where the scaling factor  $\lambda = 0.5$  controls the amount of sharpening.

#### 3.3.2 Slice-Wise Bilateral Filter

The bilateral filter weights neighboring voxels by both spatial proximity and intensity similarity. Because PET volumes often exhibit anisotropic voxel spacing, the filter is applied slice-wise. The filtered output  $V_{BF}$  is defined as:

$$V_{BF}(\mathbf{x}) = \frac{1}{W_p} \sum_{\mathbf{x}_i \in \Omega} G_{\sigma_s}(\|\mathbf{x} - \mathbf{x}_i\|) G_{\sigma_r}(|V(\mathbf{x}) - V(\mathbf{x}_i)|) V(\mathbf{x}_i) \quad (11)$$

where  $\sigma_s = 2.0$  controls the spatial kernel width and  $\sigma_r = 0.1$  controls the radiometric (intensity) similarity. The final enhanced volume is obtained by blending the residual detail layer,  $V - V_{BF}$ , back into the original image.

#### 3.3.3 Slice-Wise CLAHE

CLAHE [29] enhances local contrast by computing histograms on small grid tiles (typically  $8 \times 8$ ). To prevent noise amplification in homogeneous PET regions, each local histogram  $h(n)$  is clipped at a threshold  $\beta$ :

$$h_{clip}(n) = \begin{cases} \beta & \text{if } h(n) > \beta \\ h(n) & \text{otherwise} \end{cases} \quad (12)$$

The excess probability mass, given by  $\sum(h(n) - h_{clip}(n))$ , is redistributed uniformly across all histogram bins before computing the cumulative distribution function (CDF) used for intensity remapping.

### 3.4 Classification Experimental Setup

To validate the practical utility of the proposed enhancement, we utilized a diverse suite of modern image classification architectures, ranging from Convolutional Neural Networks (CNNs) to Vision Transformers (ViTs) and hybrid models. The specific models evaluated include:

- **Transformers & Hybrids:** Vision Transformer (ViT) [30], SwiftFormer [31].
- **Modern CNNs:** ConvNextV2 [32], EfficientNetV2 [33], GhostNetV3 [34], MobileNetV4 [35], RegNet [36].
- **State Space Models:** MambaOut [37].

### 3.4.1 Data Partitioning and Input Representation

The dataset was divided into training and validation sets using an 85/15 subject-level split to ensure that no subject contributed scans to both sets, thus preventing data leakage. For classification, inputs were constructed from 2D mid-slices extracted from both the raw and enhanced FDG-PET volumes. Performance was independently evaluated in all three anatomical planes (sagittal, coronal, and axial) to assess plane-specific sensitivity and robustness.

### 3.4.2 Training Configuration and Augmentation

All classifiers were initialized with pre-trained weights and fine-tuned using AdamW with an initial learning rate of  $2 \times 10^{-5}$  and a Cosine Annealing schedule. Cross-entropy loss was used for optimization, and half-precision (FP16) training was used to reduce memory usage. Each training slice was augmented using a standardized pipeline that included: (i) resizing to  $224 \times 224$  for ViT architectures and  $256 \times 256$  for all others, (ii) randomized blurring operations (Motion blur or Gaussian blur), (iii) geometric transformations such as affine or elastic deformation, (iv) intensity-level modifications including color jitter, gamma adjustment, and sharpening, and (v) normalization using the mean and standard deviation of the dataset. The validation set was only resized to the required model input size and normalized.

## 3.5 Quantitative Evaluation Metrics

The proposed method is quantitatively evaluated using two sets of metrics: reconstruction quality and classification.

### 3.5.1 Reconstruction Quality Metrics

Peak Signal-to-Noise Ratio (PSNR)

PSNR measures the ratio of the signal's maximum possible power to the power of corrupting noise/distortion.

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{(\max(V))^2}{\text{MSE}} \right) \quad (13)$$

where  $\text{MSE} = \frac{1}{N} \sum_{\mathbf{x}} (V_{\text{raw}}(\mathbf{x}) - V_{\text{enh}}(\mathbf{x}))^2$ .

Structural Similarity Index Measure (SSIM)

SSIM evaluates the perceived change in structural information. For two local windows  $x$  and  $y$ :

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (14)$$

where  $\mu$  denotes the mean,  $\sigma^2$  the variance and  $\sigma_{xy}$  the covariance. We emphasize that PSNR and SSIM are not interpreted as measures of fidelity to a ground-truth reference, since no high-resolution ground truth exists for in vivo FDG-PET. Instead, they serve as relative consistency metrics that quantify how strongly

an enhancement deviates from the original signal structure, complementing task-relevant measures such as Contrast-to-Noise Ratio (CNR), EPI, and NG (defined below).

Contrast-to-Noise Ratio (CNR)

Calculated between high-uptake ( $P_{80}$ ) and low-uptake ( $P_{20}$ ) regions defined by the raw image distribution:

$$\text{CNR} = \frac{\mu_{high} - \mu_{low}}{\sqrt{\sigma_{high}^2 + \sigma_{low}^2}} \quad (15)$$

Edge Preservation Index (EPI)

To assess how effectively the enhancement preserves true anatomical boundaries without introducing an artificial structure, we define the EPI. The metric is computed as the Pearson correlation between the gradient magnitudes of the raw volume ( $V_{raw}$ ) and the enhanced volume ( $V_{enh}$ ).

Let the magnitude of the gradient be  $\Delta V = \sqrt{\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2}$ . The EPI is calculated as:

$$\text{EPI} = \frac{\sum_{\mathbf{x} \in M} (\Delta V_{raw}(\mathbf{x}) - \mu_{raw})(\Delta V_{enh}(\mathbf{x}) - \mu_{enh})}{\sqrt{\sum_{\mathbf{x} \in M} (\Delta V_{raw}(\mathbf{x}) - \mu_{raw})^2} \sqrt{\sum_{\mathbf{x} \in M} (\Delta V_{enh}(\mathbf{x}) - \mu_{enh})^2}} \quad (16)$$

Here,  $\mu_{raw}$  and  $\mu_{enh}$  denote the mean gradient magnitudes within the brain mask  $M$ . Values near 1.0 indicate that structural edges are preserved and improved proportionally, while lower values suggest geometric distortion, suppression of true boundaries, or generation of spurious edges.

Noise Gain (NG).

We define noise gain as an image-domain proxy that quantifies how much additional high-frequency variance an enhancement method injects into tissue that appears homogeneous in the raw scan. The metric is therefore intended for relative comparison of enhancement methods under identical acquisition conditions, rather than as an absolute physiological noise measurement.

We first compute the magnitude of the gradient of the raw volume  $V_{raw}$ ,

$$\Delta V_{raw}(\mathbf{x}) = \sqrt{\left(\frac{\partial V_{raw}}{\partial x}\right)^2 + \left(\frac{\partial V_{raw}}{\partial y}\right)^2 + \left(\frac{\partial V_{raw}}{\partial z}\right)^2}, \quad (17)$$

and use it to identify voxels that are locally homogeneous in the raw image. As an operational approximation to “flat tissue”, we construct a mask  $\Omega_{flat}$  consisting of voxels whose raw gradient magnitude lies below the 30th percentile ( $P_{30}$ ):

$$\Omega_{flat} = \{\mathbf{x} \in M \mid \Delta V_{raw}(\mathbf{x}) \leq P_{30}(\Delta V_{raw})\} \quad (18)$$

This mask is computed once from  $V_{raw}$  and reused for all enhanced volumes, so that the region in which NG is evaluated does not depend on the enhancement method. In practice, for ADNI FDG-PET scans,  $\Omega_{flat}$  predominantly covers ventricular Cerebrospinal Fluid (CSF), deep white matter, and extra-cortical low-uptake regions, which are known to be physiologically homogeneous and dominated by Poisson noise [38].

The High-frequency residuals  $R$  are then computed by subtracting a smoothed version of the volume (which is obtained via Gaussian filtering with  $\sigma = 2.0$ ) from the original:

$$R(\mathbf{x}) = V(\mathbf{x}) - (G_{\sigma} * V)(\mathbf{x}) \quad (19)$$

We denote by  $R_{\text{raw}}$  and  $R_{\text{enh}}$  the residuals of the raw and enhanced images, respectively. NG is then defined as the ratio of the standard deviations of these residuals in the enhanced and raw images, restricted to  $\Omega_{\text{flat}}$ :

$$\text{NG} = \frac{\sigma(R_{\text{enh}}(\mathbf{x}))_{\mathbf{x} \in \Omega_{\text{flat}}}}{\sigma(R_{\text{raw}}(\mathbf{x}))_{\mathbf{x} \in \Omega_{\text{flat}}}} \quad (20)$$

Values near  $\text{NG} \approx 1.0$  indicate that noise levels are preserved, while  $\text{NG} > 1.0$  indicates noise amplification, and  $\text{NG} < 1.0$  shows noise reduction.

For structure-preserving enhancement, an NG close to unity is desirable, as it indicates that contrast is increased without altering the underlying noise statistics. Although noise reduction may be preferred in low-dose PET reconstruction, the present work focuses on enhancement rather than denoising; therefore, the target behavior is noise preservation rather than noise suppression.

## 4 Results

### 4.1 Quantitative Reconstruction Quality

We first assess the proposed Hybrid enhancement using standard image quality metrics and compare it with established baseline methods. Table 3 reports the mean  $\pm$  standard deviation of each metric computed across all  $N = 1928$  FDG-PET volumes in the dataset, where each volume contributes one measurement per metric.

**Table 3:** Comparison of image enhancement metrics (Mean  $\pm$  Std). Arrows indicate favorable direction ( $\uparrow$  higher is better,  $\downarrow$  lower is better,  $\approx 1.0$  is desirable).

Method	CNR ( $\uparrow$ )	EPI ( $\uparrow$ )	Noise gain ( $\approx 1.0$ )	PSNR ( $\uparrow$ )	SSIM ( $\uparrow$ )
Raw Baseline	$5.26 \pm 0.32$	1.00 (Ref)	1.00 (Ref)	- (Ref)	1.00 (Ref)
Laplacian	$5.26 \pm 0.32$	$0.994 \pm 0.002$	$0.99 \pm 0.01$	$39.29 \pm 1.73$	$0.999 \pm 0.000$
DoG	$5.28 \pm 0.34$	$0.962 \pm 0.006$	$1.06 \pm 0.05$	$28.06 \pm 1.01$	$0.988 \pm 0.002$
Unsharp	$5.13 \pm 0.33$	$0.993 \pm 0.001$	$1.04 \pm 0.02$	$37.11 \pm 1.46$	$0.998 \pm 0.000$
Bilateral	$5.40 \pm 0.26$	$0.987 \pm 0.004$	$1.28 \pm 0.04$	$34.49 \pm 0.42$	$0.993 \pm 0.000$
CLAHE	$5.21 \pm 0.33$	$0.954 \pm 0.012$	$1.42 \pm 0.06$	$24.19 \pm 0.91$	$0.979 \pm 0.002$
<b>Hybrid (Ours)</b>	$5.28 \pm 0.33$	$0.981 \pm 0.002$	$1.01 \pm 0.03$	$32.06 \pm 1.08$	$0.996 \pm 0.001$

*Note:* For the Raw baseline, PSNR and SSIM are defined as identity references (PSNR and SSIM relative to themselves are not meaningful); therefore, they are indicated as “Reference” rather than omitted. High SSIM values for the Laplacian variant indicate that the operator preserves geometric structure and alters only local contrast, resulting in minimal change in structural similarity.

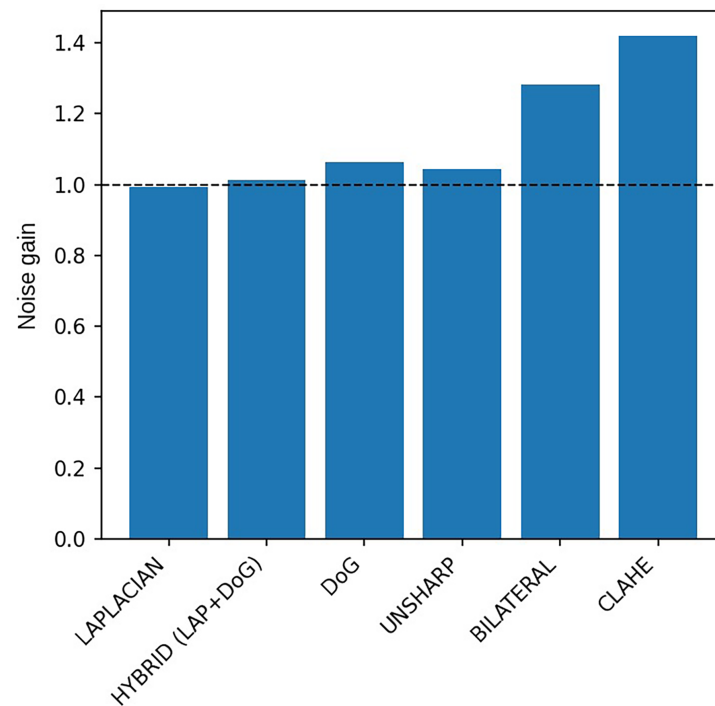
The analysis highlights a common trade-off in existing enhancement techniques between increased contrast and loss of signal fidelity. The Bilateral filter achieves the highest apparent contrast gain (CNR = 5.40), but this improvement is accompanied by substantial noise amplification ( $\text{NG} = 1.28$ ) and a reduced SSIM relative to the Laplacian baseline, suggesting that texture is being overly smoothed to elevate regional averages. CLAHE similarly boosts contrast aggressively, yet produces the highest noise amplification ( $\text{NG} = 1.42$ ) and the lowest structural similarity (SSIM = 0.979) among the evaluated methods.

In contrast, the proposed Hybrid method exhibits a more balanced performance profile. It achieves a CNR of 5.28, exceeding the Raw baseline (5.26) and unsharp masking (5.13), while maintaining a Noise Gain of 1.01. This near-unity NG value is particularly important: it indicates that the method enhances genuine

metabolic gradients without modifying the underlying noise distribution, preserving radiomic stability, and ensuring that quantitative analyses remain reliable.

#### 4.1.1 Analysis of Enhancement Efficiency

To further characterize the relationship between sharpening strength and artifact formation, we examined noise behavior as a function of contrast gain. Fig. 2 summarizes the Noise Gain values for all methods. Both Bilateral filtering and CLAHE substantially increase the noise floor ( $NG > 1.02$ ), reflecting their tendency to amplify high-frequency variance. In contrast, the proposed Hybrid method remains effectively anchored to the baseline noise level, reinforcing its ability to enhance contrast without introducing spurious detail.

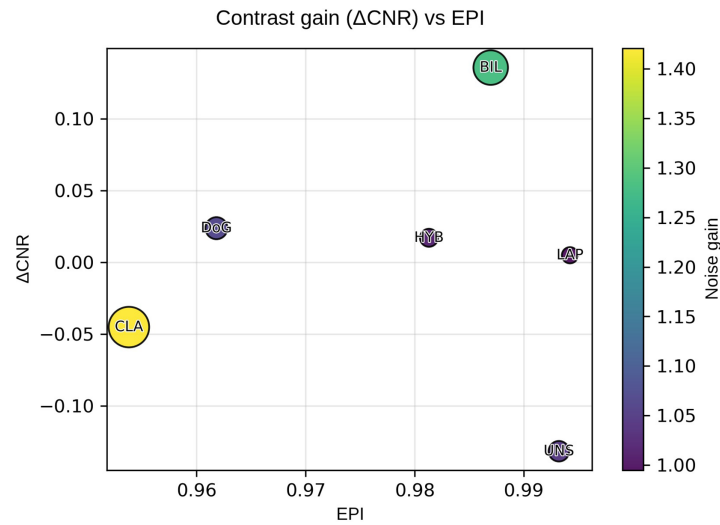


**Figure 2:** Noise gain comparison. The dashed line represents the noise level of the raw input. The Hybrid method introduces negligible noise amplification compared to Bilateral or CLAHE.

Fig. 3 illustrates the Enhancement Efficiency by plotting the change in contrast ( $\Delta\text{CNR}$ ) against the EPI. Bubble size and color encode the Noise Gain, with smaller, purple markers indicating better noise behavior. In this representation, an ideal enhancement method would appear in the upper-right quadrant with a minimal bubble radius.

CLAHE and Unsharp Masking fall in the negative  $\Delta\text{CNR}$  region, indicating that their sharpening reduces the overall contrast-to-noise ratio. Unsharp Masking preserves edges well ( $\text{EPI} > 0.99$ ), but its high noise amplification offsets this benefit. CLAHE performs worse, showing reduced edge preservation ( $\text{EPI} < 0.96$ ) and a large noise footprint.

Bilateral filtering provides the strongest contrast gain ( $\Delta\text{CNR} \approx 0.13$ ), but the improvement is accompanied by substantial noise amplification ( $NG \approx 1.28$ ). This suggests that although regional contrast is increased, the underlying texture statistics are altered enough that the enhanced image no longer faithfully represents the original tracer distribution.



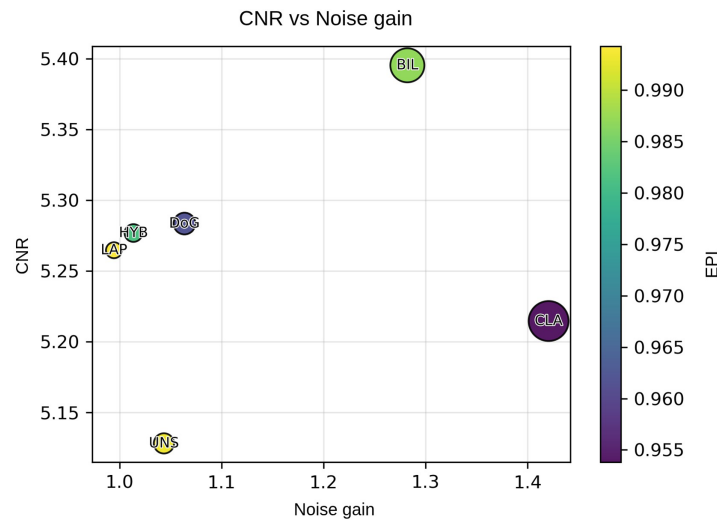
**Figure 3:** Enhancement efficiency:  $\Delta\text{CNR}$  vs. EPI. The bubble size represents Noise Gain. The Hybrid method (HYB) provides positive contrast gain while maintaining a small noise gain.

DoG yields moderate contrast improvements, but exhibits reduced edge preservation (EPI  $\approx$  0.96).

Laplacian filtering preserves edges more reliably (high EPI, minimal noise), but contributes little to contrast enhancement ( $\Delta\text{CNR} \approx 0$ ).

The Hybrid method balances these properties. It achieves the positive contrast gain of DoG while shifting rightward in edge preservation (EPI  $\approx$  0.982), and maintains a near-unity NG comparable to the Laplacian. This combination indicates that the Hybrid method enhances meaningful structure without incurring the noise penalties seen in Bilateral or DoG methods.

Fig. 4 shows the relationship between CNR and NG. An ideal enhancement method would appear in the upper-left region of the plot (high CNR with minimal noise amplification).



**Figure 4:** The  $y$ -axis denotes contrast (higher is better), while the  $x$ -axis reflects Noise Gain (lower is better). Bubble size encodes the magnitude of Noise Gain, and color represents Edge Preservation (yellow indicating high fidelity, purple indicating reduced fidelity). The Hybrid method (HYB) increases contrast vertically while remaining on the left side of the plot, indicating minimal noise amplification.

The Bilateral filter achieves the highest CNR (5.40) but does so at the cost of substantial noise amplification ( $NG \approx 1.28$ ), as indicated by its far-right position and large marker size. CLAHE performs even less favorably: it exhibits elevated noise levels ( $NG > 1.4$ ), a comparatively low CNR, and reduced edge integrity (darker color), placing it at the bottom-right of the plot.

Unsharp Masking shows stable edge preservation (light color) but produces limited contrast improvement, remaining near the lower end of the CNR axis ( $< 5.15$ ), suggesting that, while structurally safe, it does not significantly enhance metabolic contrast relative to the raw input.

The most informative comparison involves the Laplacian (LAP), DoG, and Hybrid (HYB) methods. Laplacian filtering offers the most stable noise behavior ( $NG < 1.0$ ) and strong edge fidelity, but yields minimal improvement in CNR.

DoG increases CNR but shifts rightward ( $NG \approx 1.06$ ) and shows reduced edge preservation, indicating a trade-off between added contrast and structural accuracy.

The Hybrid method effectively balances these factors. It achieves a CNR comparable to DoG ( $\approx 5.28$ ) while maintaining a noise profile close to the Laplacian ( $NG \approx 1.01$ ). This placement along the Pareto frontier demonstrates that the Hybrid approach enhances contrast without incurring significant noise penalties.

#### 4.1.2 Sensitivity Analysis of Modulation Parameters

To assess the robustness of the proposed enhancement framework with respect to its modulation parameters, we conducted a lightweight sensitivity analysis focusing on the three parameters explicitly controlling enhancement strength: the Laplacian contrast gain  $k$ , the DoG mixing factor  $\alpha$ , and the gradient gate threshold  $\tau$ . These parameters shape the amplitude of contrast amplification and detail injection, but do not affect the spatial support of the underlying filters. The sensitivity analysis results are listed in [Table 4](#).

**Table 4:** Sensitivity analysis of the hybrid enhancement framework with respect to the modulation parameters  $k$ ,  $\alpha$ , and  $\tau$ .  $\Delta\text{CNR}$  denotes  $\text{CNR}_{\text{enh}} - \text{CNR}_{\text{raw}}$ .  $NG$  values close to 1 indicate noise preservation.

$k$	$\alpha$	$\tau$	$\Delta\text{CNR}$	EPI	NG
1.50	0.40	0.005	$+0.017 \pm 0.017$	$0.989 \pm 0.001$	$1.003 \pm 0.019$
1.50	0.40	0.010	$+0.017 \pm 0.017$	$0.989 \pm 0.001$	$0.998 \pm 0.019$
1.50	0.40	0.020	$+0.014 \pm 0.016$	$0.989 \pm 0.001$	$0.973 \pm 0.021$
1.50	0.60	0.005	$+0.021 \pm 0.023$	$0.983 \pm 0.002$	$1.012 \pm 0.023$
1.50	0.60	0.010	$+0.021 \pm 0.022$	$0.984 \pm 0.002$	$1.005 \pm 0.023$
1.50	0.60	0.020	$+0.017 \pm 0.022$	$0.984 \pm 0.002$	$0.967 \pm 0.027$
1.50	0.70	0.005	$+0.022 \pm 0.025$	$0.980 \pm 0.002$	$1.017 \pm 0.025$
1.50	0.70	0.010	$+0.022 \pm 0.025$	$0.980 \pm 0.002$	$1.010 \pm 0.026$
1.50	0.70	0.020	$+0.018 \pm 0.024$	$0.981 \pm 0.002$	$0.966 \pm 0.029$
2.00	0.40	0.005	$+0.016 \pm 0.019$	$0.987 \pm 0.001$	$1.005 \pm 0.021$
2.00	0.40	0.010	$+0.016 \pm 0.019$	$0.987 \pm 0.001$	$1.000 \pm 0.021$
2.00	0.40	0.020	$+0.014 \pm 0.018$	$0.987 \pm 0.001$	$0.974 \pm 0.023$
2.00	0.60	0.005	$+0.020 \pm 0.024$	$0.981 \pm 0.002$	$1.014 \pm 0.025$
2.00	0.60	0.010	$+0.020 \pm 0.024$	$0.981 \pm 0.002$	$1.008 \pm 0.025$
2.00	0.60	0.020	$+0.016 \pm 0.023$	$0.982 \pm 0.002$	$0.970 \pm 0.029$
2.00	0.70	0.005	$+0.021 \pm 0.027$	$0.978 \pm 0.002$	$1.020 \pm 0.027$
2.00	0.70	0.010	$+0.021 \pm 0.027$	$0.978 \pm 0.002$	$1.013 \pm 0.027$

(Continued)

**Table 4 (continued)**

$k$	$\alpha$	$\tau$	$\Delta\text{CNR}$	EPI	NG
2.00	0.70	0.020	$+0.017 \pm 0.026$	$0.978 \pm 0.002$	$0.970 \pm 0.031$
2.50	0.40	0.005	$+0.015 \pm 0.021$	$0.985 \pm 0.002$	$1.007 \pm 0.023$
2.50	0.40	0.010	$+0.015 \pm 0.021$	$0.985 \pm 0.002$	$1.002 \pm 0.024$
2.50	0.40	0.020	$+0.013 \pm 0.020$	$0.985 \pm 0.002$	$0.977 \pm 0.026$
2.50	0.60	0.005	$+0.018 \pm 0.026$	$0.979 \pm 0.002$	$1.018 \pm 0.027$
2.50	0.60	0.010	$+0.018 \pm 0.026$	$0.979 \pm 0.002$	$1.011 \pm 0.028$
2.50	0.60	0.020	$+0.015 \pm 0.025$	$0.979 \pm 0.002$	$0.974 \pm 0.031$
2.50	0.70	0.005	$+0.019 \pm 0.029$	$0.976 \pm 0.002$	$1.024 \pm 0.029$
2.50	0.70	0.010	$+0.019 \pm 0.029$	$0.976 \pm 0.002$	$1.017 \pm 0.029$
2.50	0.70	0.020	$+0.015 \pm 0.028$	$0.976 \pm 0.002$	$0.974 \pm 0.033$

We evaluated combinations of  $k \in \{1.5, 2.0, 2.5\}$ ,  $\alpha \in \{0.4, 0.6, 0.7\}$ , and  $\tau \in \{0.005, 0.01, 0.02\}$ . For each configuration, we computed reconstruction quality metrics (CNR, EPI) and noise gain (NG), without retraining downstream classifiers.

Across all tested configurations, the proposed method exhibited stable behavior. Contrast-to-noise improvements remained consistently positive ( $\Delta\text{CNR} \approx 0.013\text{--}0.022$ ), while structural fidelity was preserved with high edge preservation indices (EPI  $\approx 0.976\text{--}0.989$ ). Noise gain remained close to unity for  $\tau \in \{0.005, 0.01\}$ , indicating that contrast enhancement did not introduce measurable noise amplification. Larger values of  $\tau$  resulted in slightly reduced contrast gain and  $\text{NG} < 1$ , reflecting a more conservative suppression of detail injection in low-gradient regions.

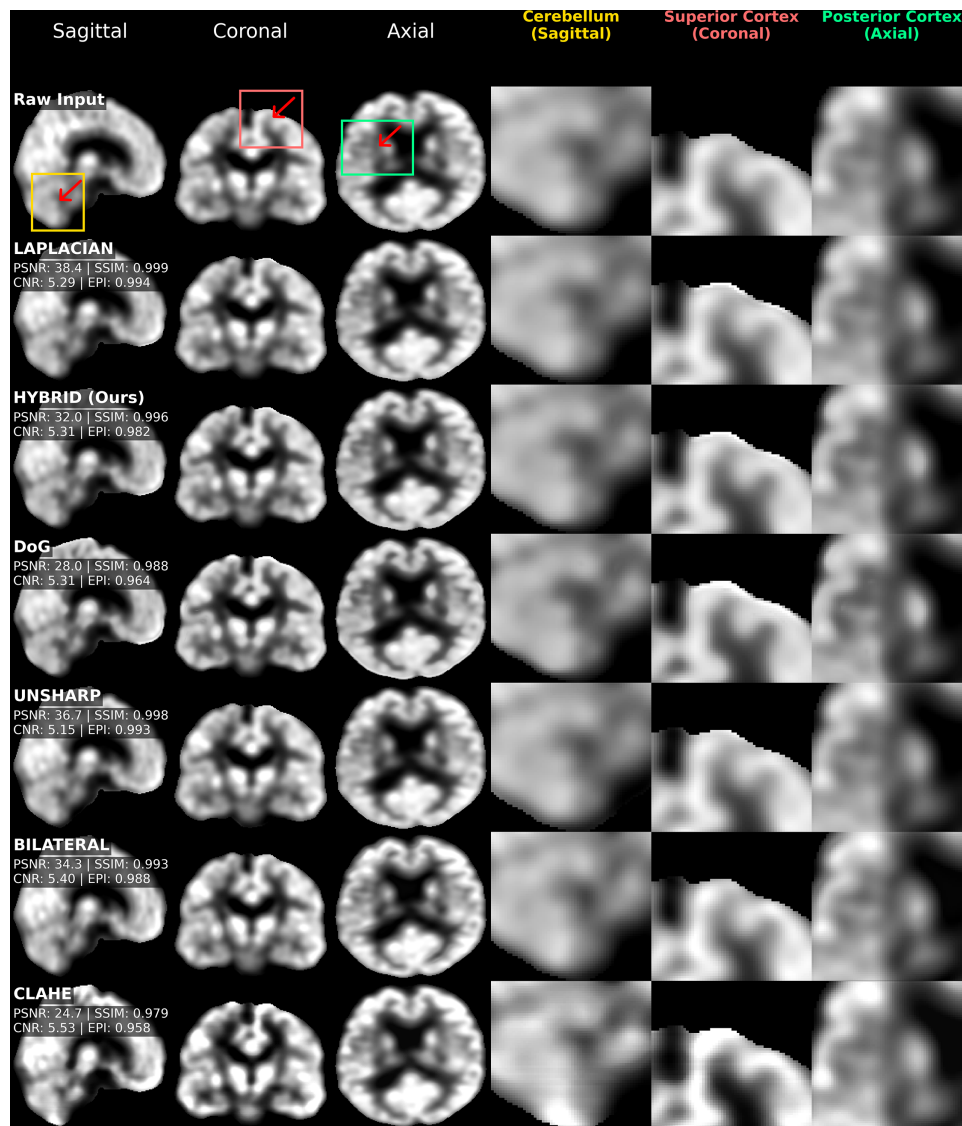
These results confirm that the method is not sensitive to precise tuning of  $k$ ,  $\alpha$ , or  $\tau$  within a physiologically reasonable range. The default configuration ( $k = 2.0$ ,  $\alpha = 0.6$ ,  $\tau = 0.01$ ) lies within the most stable region of the parameter space and provides a balanced trade-off between contrast enhancement, structural preservation, and noise stability.

#### 4.2 Qualitative Visual Assessment

Visual inspection of the enhanced volumes (Fig. 5) supports the quantitative findings and highlights characteristic behaviors of each method. The red arrows mark the regions where these differences are particularly evident.

In both the cortical ribbon (axial view) and the cerebellum (sagittal view), the Hybrid method improves structural definition by tightening the metabolic boundaries between gray and white matter. The enhancement recovers details lost to partial volume effects while preserving the underlying texture, resulting in a more coherent anatomical appearance than the Raw input.

CLAHE and Bilateral filtering show visible saturation in high-uptake regions (coronal view). Bilateral filtering produces a smoothed, “plastic” texture in which local variations are flattened, whereas CLAHE introduces block-like high-intensity artifacts that obscure subtle metabolic gradients. Both behaviors are consistent with their elevated Noise Gain and reduced SSIM.



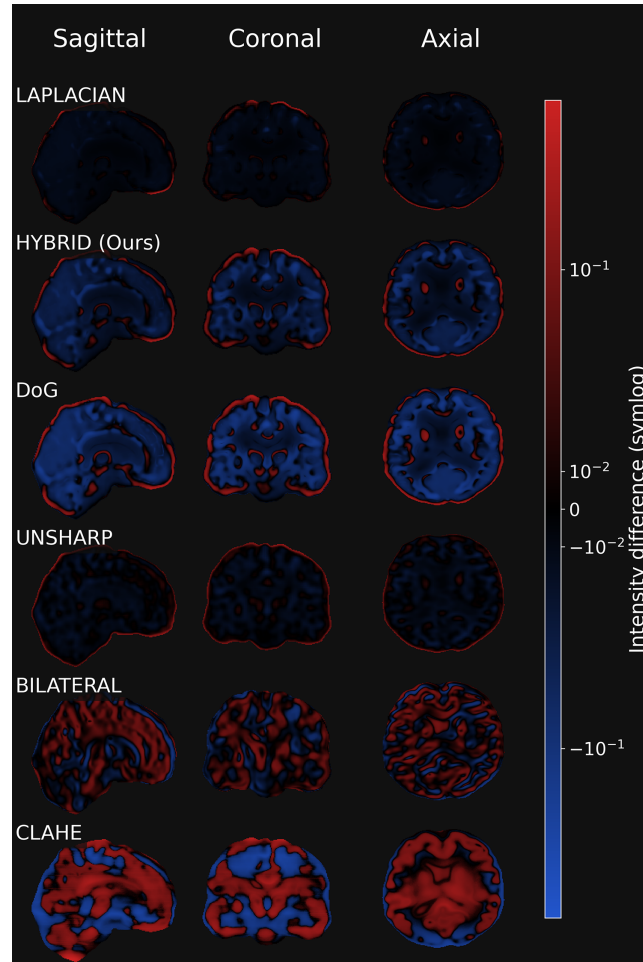
**Figure 5:** Montage comparison on axial, coronal, and sagittal planes. Note the clearer definition of the cortical ribbon in the Hybrid method compared to Raw.

The DoG method yields the sharpest apparent edges, but the enhancement is excessively strong. The resulting boundaries appear etched and exaggerated relative to the true tracer distribution. This aligns with its lower Edge Preservation Index (EPI = 0.964), indicating that the method modifies the edge geometry rather than simply improving the existing structure.

Therefore, the visual assessments indicate that the proposed Hybrid method combines the structural fidelity of the Laplacian with the contrast enhancement capability of DoG. The resulting images preserve the natural appearance of the Raw scan, avoiding the artificial texture introduced by Bilateral filtering or CLAHE, while providing sufficient contrast improvement to reveal subtle metabolic features.

#### 4.2.1 Structural Decomposition Analysis

Difference mapping provides additional information on how each enhancement method modifies the underlying signal by visualizing where the intensity has been added or removed (see Fig. 6).



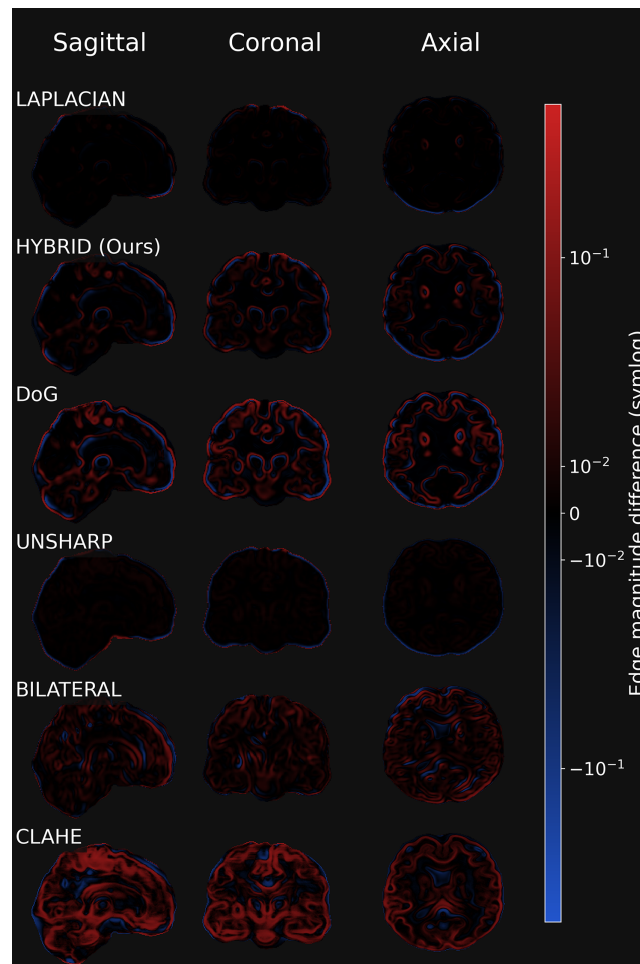
**Figure 6:** Intensity difference ( $\Delta V$ ). Note the massive, non-anatomical intensity shifts in CLAHE compared to the localized enhancement of the Hybrid method.

CLAHE produces large low-frequency shifts across wide regions of the brain, visible as broad red and blue areas. The global intensity changes suggest that the method alters the radiotracer distribution rather than improving existing structure, raising concerns about potential errors in quantitative SUV measurements.

The Bilateral filter generates a mottled pattern of localized intensity changes, indicating that it modifies fine-scale textures in a way that disrupts the natural appearance of the scan, which is consistent with the “plastic” visual effect observed in the qualitative analysis.

The Hybrid method reveals localized, structurally significant changes, predominantly confined to the cortical ribbon. It avoids the widespread background shifts of CLAHE and the texture perturbations of Bilateral filtering, altering intensity only where necessary to recover anatomical boundaries.

The edge-magnitude difference maps (Fig. 7) further illustrate how each method alters structural boundaries.



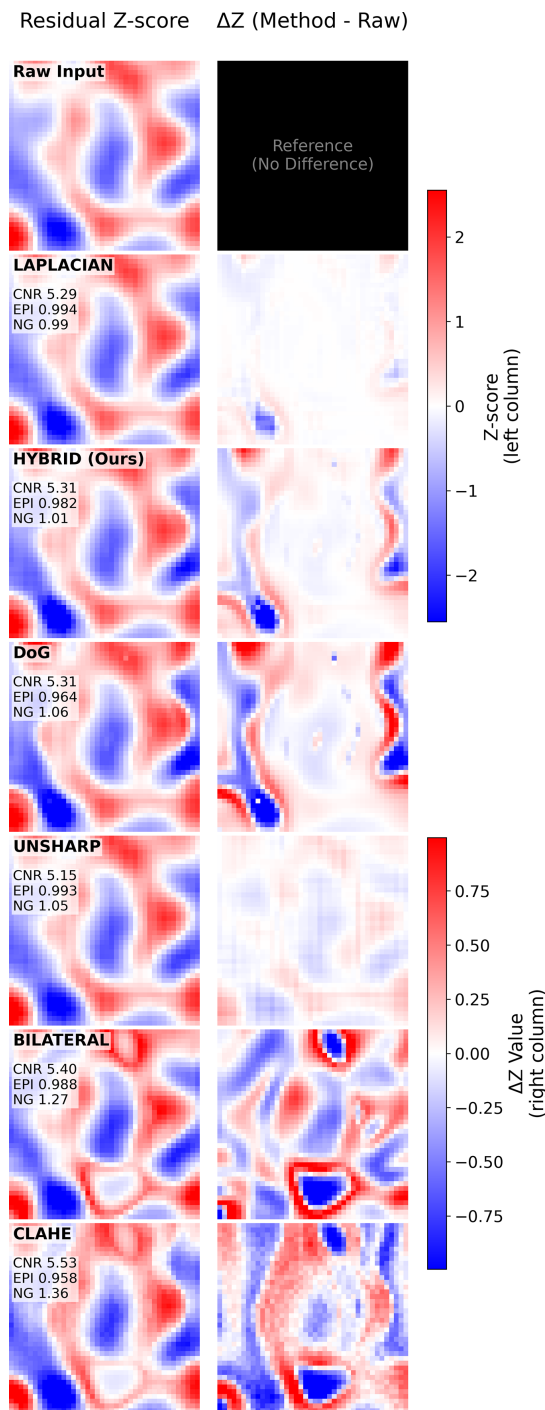
**Figure 7:** Edge magnitude difference ( $\Delta|\nabla V|$ ). The Hybrid method produces sharp, thin edge enhancements (selective), whereas DoG thickens edges globally.

The DoG method (row 3) produces widespread red contours throughout the volume, indicating broad gradient amplification. This uniform strengthening of edges thickens boundaries indiscriminately and can cause adjacent metabolic regions to merge, thereby exacerbating partial-volume effects.

The Hybrid method (row 2) produces thinner, more anatomically consistent edge enhancements. Sharpening is concentrated along true cortical boundaries (gyri and sulci) and suppressed in adjacent white matter regions. The behavior confirms that the gradient-weighting term  $W_{grad}$  effectively limits sharpening to meaningful anatomical transitions, avoiding the over-etched appearance seen in the DoG output.

The residual noise analysis (Fig. 8) provides a detailed view of how each algorithm alters the image's high-frequency content. The rightmost column,  $\Delta Z$  (Method–Raw), isolates the specific residual components introduced by each enhancement method.

The  $\Delta Z$  maps for Bilateral filtering and CLAHE show prominent red and blue patches, indicating substantial local intensity rearrangements. Bilateral filtering produces artificial plateaus, disrupting the natural stochastic texture of PET data, while CLAHE introduces block-like artifacts due to its aggressive histogram redistribution. The patterns align with their elevated Noise Gain values ( $NG > 1.25$ ) and suggest that both methods may compromise the stability of the radiomic feature.



**Figure 8:** Noise residual analysis ( $\Delta Z$ ). Residual Z-score for each method (left column).  $\Delta Z = Z_{enh} - Z_{raw}$ , isolating changes introduced by enhancement (right column). The two columns use independent color scales (indicated by labels). The Hybrid method's  $\Delta Z$  shows structurally coherent residuals rather than random noise speckle, indicating true detail recovery.

DoG and Unsharp Masking display speckled noise patterns that extend into homogeneous tissue, which confirms that linear sharpening approaches amplify background noise and structure equally when no gating mechanism is applied, resulting in a reduced signal-to-noise ratio in regions such as white matter.

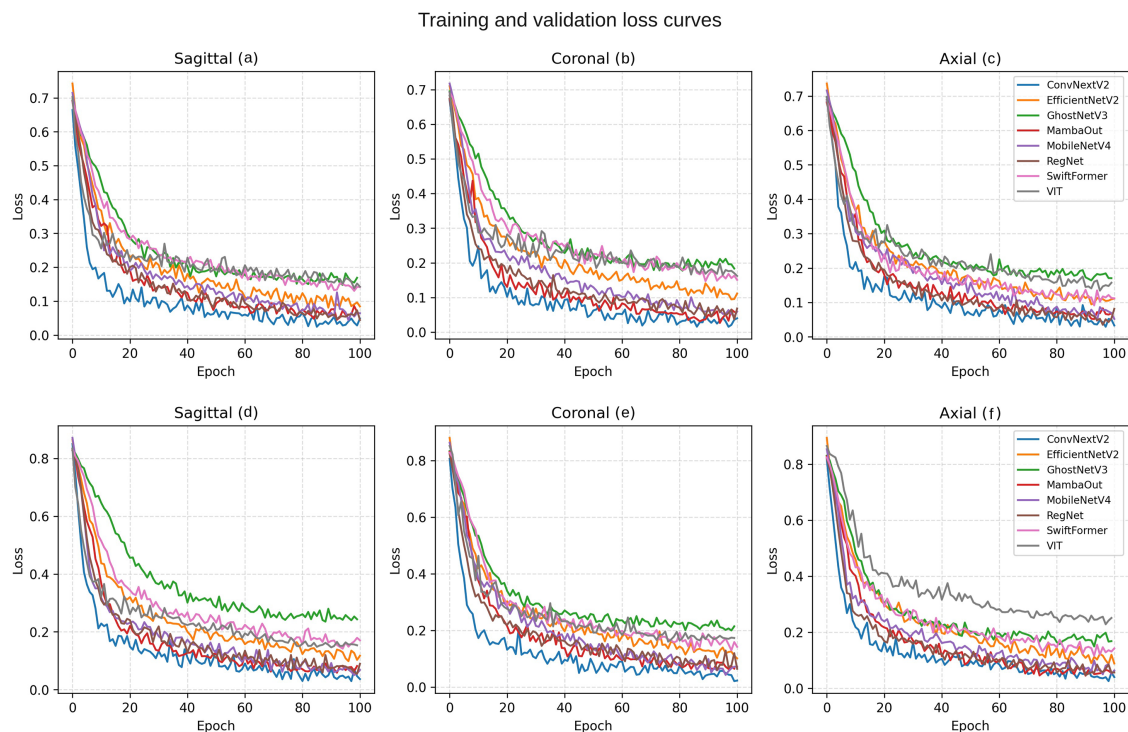
The Laplacian baseline exhibits minimal changes in its  $\Delta Z$  map, reflecting its low Noise Gain ( $NG \approx 0.99$ ). Although this indicates high stability, it also highlights the method's limited ability to recover subtle structural information in the absence of an explicit detail-injection mechanism.

The Hybrid method displays sparse yet anatomically aligned residuals. Deviations occur primarily along cortical boundaries rather than in homogeneous tissue, distinguishing it from the speckled patterns of DoG or the patch-like artifacts of Bilateral filtering. This pattern confirms the intended behavior of the gradient-gating mechanism: noise remains largely unchanged in flat regions, and detail is injected only at meaningful anatomical transitions, effectively restoring structure lost to partial-volume effects without introducing artificial texture.

### 4.3 MCI Classification Performance

To evaluate the diagnostic performance of classifiers trained on enhanced PET images, we adopt standard binary classification metrics derived from the confusion matrix: Accuracy (ACC), Sensitivity (SEN), Specificity (SPE), Matthews Correlation Coefficient (MCC), Area Under Curve (AUC), Cohen's Kappa (CK), and F1 Score. To quantify variability, all models were trained with multiple random seeds, and results are reported as mean  $\pm$  standard deviation.

All classification models were trained using identical data splits, optimization settings, and early-stopping criteria to ensure a fair and controlled comparison. Across all architectures and anatomical planes, training and validation loss curves showed stable convergence, with no signs of divergence or overfitting. For illustration, Fig. 9 presents the training and validation loss curves for MCI classification using the enhanced PET volumes. Training curves for the raw-input setting exhibited the same qualitative convergence behavior and are therefore omitted for brevity.



**Figure 9:** Training and validation loss curves for MCI classification with enhanced PET data. Top row: training loss for (a) sagittal, (b) coronal, and (c) axial planes. Bottom row: corresponding validation loss for (d) sagittal, (e) coronal, and (f) axial planes.

Accuracy, sensitivity, specificity, and AUC trends were consistent across classes and models, supporting the robustness of the reported performance gains.

The proposed Hybrid enhancement yields consistent improvements in MCI classification across all architectures and anatomical planes evaluated (Fig. 10). Among the three views, the axial plane provided the strongest diagnostic performance, both before and after enhancement.

In the axial plane, MobileNetV4 achieved the highest overall performance when trained on enhanced images, reaching a mean accuracy of  $0.96 \pm 0.05$  with an AUC of  $0.99 \pm 0.07$ . This represents a notable improvement over the raw baseline ( $ACC = 0.93 \pm 0.05$ ,  $AUC = 0.96 \pm 0.08$ ), resulting in a reduction in the mean misclassification rate from approximately 7% to 4%.

For MambaOut in the axial plane, mean sensitivity improved from  $0.89 \pm 0.07$  (raw) to  $0.94 \pm 0.08$  (enhanced), indicating that the Hybrid method helps recover subtle hypometabolic patterns associated with early cognitive decline that may be obscured in raw PET images.

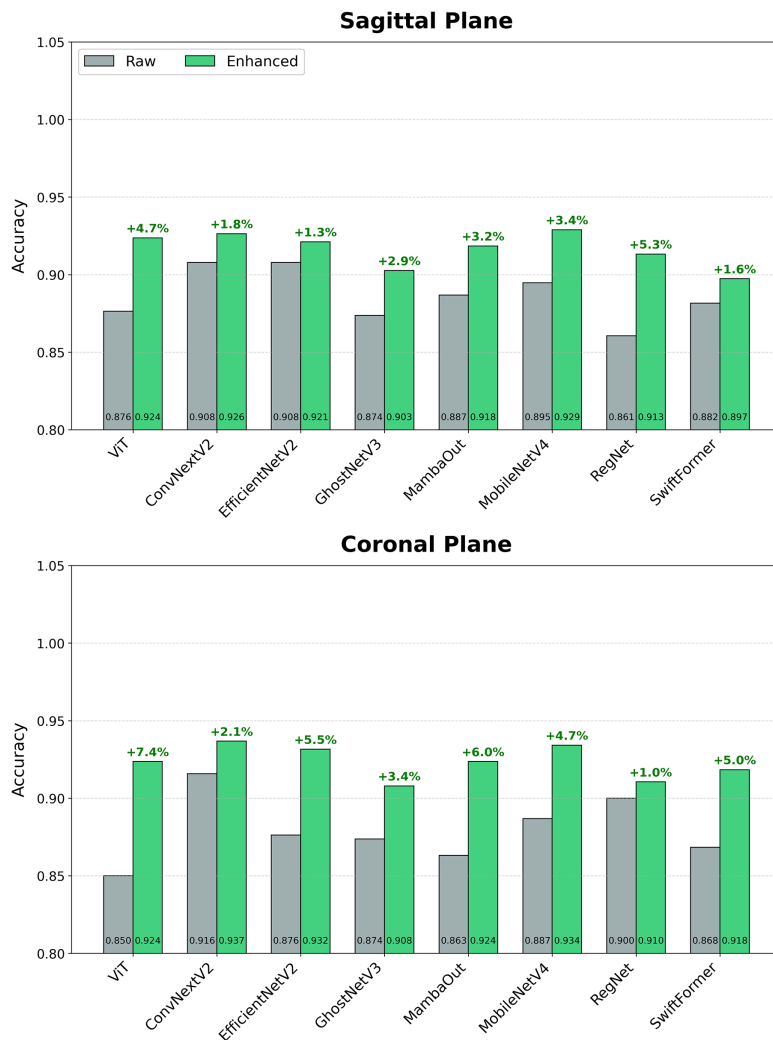
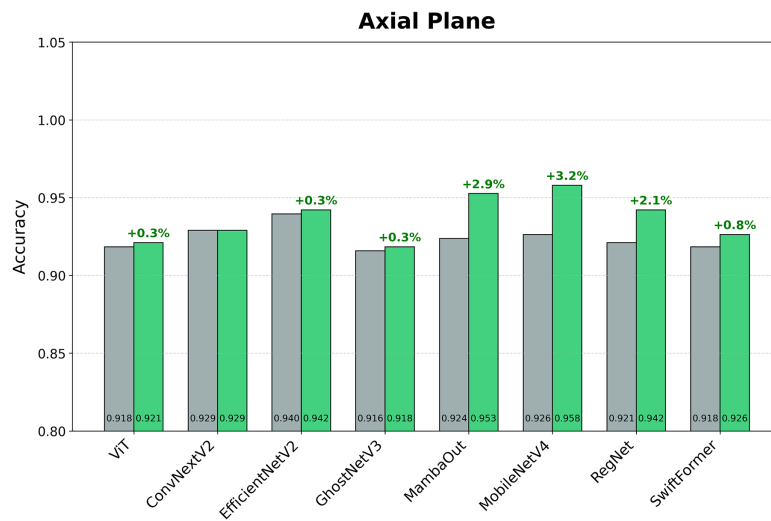
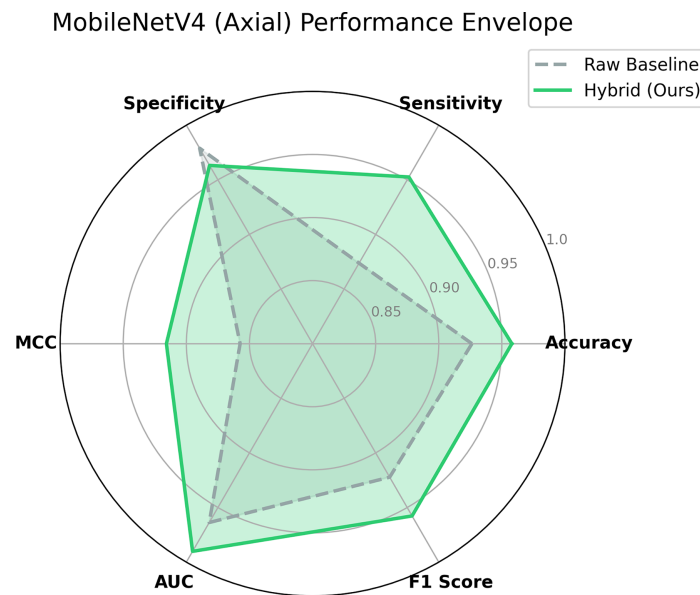


Figure 10: (Continued)



**Figure 10:** Impact of enhancement on classification accuracy across anatomical planes. The proposed method (green) consistently outperforms the raw baseline (grey) across all tested architectures. The + symbol denotes the relative change with respect to the corresponding baseline (unenhanced) input. Exact numerical values corresponding to each bar are reported in [Tables 5](#) and [6](#).

The improvement extends beyond accuracy. The radar plot in [Fig. 11](#) shows that enhancement broadens the performance envelope across all evaluation metrics for MobileNetV4. Sensitivity, which is critical for early detection, showed notable gains. Although multiple architectures were evaluated, MobileNetV4 is shown as a representative case in [Fig. 11](#) because it consistently achieved the strongest and most stable performance gains across enhancement settings and anatomical planes. Similar trends were observed for other models and are reported in [Tables 5](#) and [6](#) and [Fig. 10](#).



**Figure 11:** Holistic performance comparison for the best model (MobileNetV4, Axial). MobileNetV4 achieved the most consistent and highest overall gains across architectures and anatomical planes; therefore, it is used here as a representative example to summarize the effect of the proposed enhancement across evaluation metrics. Exact numerical values are reported in [Tables 5](#) and [6](#).

All classification metrics for all models are listed in [Tables 5 and 6](#).

The results on raw data reveal clear differences across anatomical planes and model families. The axial plane consistently provides the strongest diagnostic performance, with EfficientNetV2 and ConvNextV2 achieving mean accuracies of  $0.94 \pm 0.08$  and  $0.93 \pm 0.04$ , respectively, and AUC values of 0.99 and 0.97. These high axial scores reflect the fact that early hypometabolism patterns characteristic of MCI are most prominent and spatially coherent in this orientation. In contrast, sagittal and coronal views show broader variability and generally reduced sensitivity, suggesting a weaker representation of subtle metabolic deficits. Across architectures, performance also varies: transformer-based ViT tends to underperform relative to modern convolutional models, while lightweight networks such as GhostNetV3 and SwiftFormer show lower MCC and Kappa scores, indicating greater susceptibility to noise-driven misclassifications.

Applying the proposed Hybrid enhancement produces consistent improvements across all metrics, architectures, and planes. The most visible gains appear in sensitivity and MCC, both of which indicate that classifiers detect a greater proportion of true MCI cases and make more reliable decisions. In the sagittal plane, ViT accuracy increases from  $0.88 \pm 0.04$  to  $0.92 \pm 0.04$ , while MCC improves from  $0.75 \pm 0.06$  to  $0.85 \pm 0.08$ . The coronal plane, previously the weakest performer, demonstrates marked improvements across all models: ConvNextV2 increases from  $0.92 \pm 0.04$  to  $0.94 \pm 0.03$ , and EfficientNetV2 reaches  $0.93 \pm 0.03$  with an AUC of  $0.96 \pm 0.08$ . The axial plane continues to yield the strongest results, with MobileNetV4 achieving the highest overall performance ( $ACC = 0.96 \pm 0.05$ ,  $AUC = 0.99 \pm 0.07$ ), accompanied by corresponding increases in F1 and Kappa that reflect a reduction in both false positives and false negatives.

**Table 5:** Quantitative evaluation of the proposed FDG-PET enhancement method in MCI classification task. Evaluated on Raw data. Values are reported as Mean  $\pm$  Std.

Plane	Model	ACC	SEN	SPE	MCC	AUC	CK	F1
Sagittal	ViT	$0.88 \pm 0.04$	$0.87 \pm 0.04$	$0.88 \pm 0.03$	$0.75 \pm 0.06$	$0.92 \pm 0.02$	$0.73 \pm 0.05$	$0.88 \pm 0.03$
	ConvNextV2	$0.91 \pm 0.05$	$0.89 \pm 0.02$	$0.93 \pm 0.02$	$0.82 \pm 0.05$	$0.95 \pm 0.05$	$0.85 \pm 0.06$	$0.91 \pm 0.03$
	EfficientNetV2	$0.91 \pm 0.06$	$0.93 \pm 0.11$	$0.89 \pm 0.02$	$0.82 \pm 0.04$	$0.97 \pm 0.05$	$0.83 \pm 0.05$	$0.91 \pm 0.03$
	GhostNetV3	$0.87 \pm 0.03$	$0.86 \pm 0.06$	$0.89 \pm 0.07$	$0.75 \pm 0.12$	$0.92 \pm 0.06$	$0.79 \pm 0.05$	$0.87 \pm 0.07$
	MambaOut	$0.89 \pm 0.05$	$0.89 \pm 0.02$	$0.88 \pm 0.04$	$0.77 \pm 0.04$	$0.95 \pm 0.03$	$0.75 \pm 0.04$	$0.89 \pm 0.03$
	MobileNetV4	$0.89 \pm 0.10$	$0.91 \pm 0.06$	$0.88 \pm 0.10$	$0.79 \pm 0.05$	$0.95 \pm 0.02$	$0.81 \pm 0.06$	$0.90 \pm 0.05$
	RegNet	$0.86 \pm 0.09$	$0.87 \pm 0.01$	$0.85 \pm 0.02$	$0.72 \pm 0.08$	$0.95 \pm 0.04$	$0.75 \pm 0.02$	$0.86 \pm 0.08$
	SwiftFormer	$0.88 \pm 0.03$	$0.85 \pm 0.05$	$0.91 \pm 0.04$	$0.76 \pm 0.05$	$0.95 \pm 0.08$	$0.77 \pm 0.04$	$0.88 \pm 0.03$
Coronal	ViT	$0.85 \pm 0.04$	$0.78 \pm 0.03$	$0.92 \pm 0.02$	$0.71 \pm 0.09$	$0.93 \pm 0.09$	$0.74 \pm 0.07$	$0.84 \pm 0.07$
	ConvNextV2	$0.92 \pm 0.04$	$0.92 \pm 0.05$	$0.91 \pm 0.06$	$0.83 \pm 0.02$	$0.95 \pm 0.02$	$0.81 \pm 0.03$	$0.92 \pm 0.03$
	EfficientNetV2	$0.88 \pm 0.02$	$0.87 \pm 0.08$	$0.88 \pm 0.08$	$0.75 \pm 0.07$	$0.95 \pm 0.07$	$0.77 \pm 0.03$	$0.88 \pm 0.06$
	GhostNetV3	$0.87 \pm 0.01$	$0.85 \pm 0.10$	$0.90 \pm 0.07$	$0.75 \pm 0.05$	$0.92 \pm 0.07$	$0.72 \pm 0.03$	$0.87 \pm 0.08$
	MambaOut	$0.86 \pm 0.01$	$0.85 \pm 0.03$	$0.88 \pm 0.04$	$0.73 \pm 0.03$	$0.92 \pm 0.07$	$0.75 \pm 0.03$	$0.86 \pm 0.09$
	MobileNetV4	$0.89 \pm 0.05$	$0.85 \pm 0.01$	$0.92 \pm 0.10$	$0.78 \pm 0.06$	$0.94 \pm 0.03$	$0.80 \pm 0.07$	$0.88 \pm 0.06$
	RegNet	$0.90 \pm 0.06$	$0.92 \pm 0.06$	$0.88 \pm 0.04$	$0.80 \pm 0.07$	$0.94 \pm 0.10$	$0.76 \pm 0.08$	$0.90 \pm 0.06$
	SwiftFormer	$0.87 \pm 0.10$	$0.87 \pm 0.04$	$0.86 \pm 0.06$	$0.74 \pm 0.02$	$0.93 \pm 0.03$	$0.71 \pm 0.03$	$0.87 \pm 0.05$

(Continued)

**Table 5 (continued)**

Plane	Model	ACC	SEN	SPE	MCC	AUC	CK	F1
Axial	ViT	0.92 ± 0.05	0.87 ± 0.06	0.96 ± 0.04	0.84 ± 0.05	0.95 ± 0.06	0.85 ± 0.03	0.91 ± 0.03
	ConvNextV2	0.93 ± 0.04	0.93 ± 0.04	0.93 ± 0.04	0.86 ± 0.06	0.97 ± 0.08	0.83 ± 0.03	0.93 ± 0.02
	EfficientNetV2	0.94 ± 0.08	0.95 ± 0.05	0.93 ± 0.04	0.88 ± 0.10	0.99 ± 0.02	0.90 ± 0.04	0.94 ± 0.04
	GhostNetV3	0.92 ± 0.10	0.89 ± 0.05	0.94 ± 0.05	0.83 ± 0.05	0.95 ± 0.05	0.86 ± 0.02	0.91 ± 0.02
	MambaOut	0.92 ± 0.04	0.89 ± 0.07	0.95 ± 0.05	0.85 ± 0.06	0.96 ± 0.04	0.85 ± 0.03	0.92 ± 0.02
	MobileNetV4	0.93 ± 0.05	0.87 ± 0.01	0.98 ± 0.12	0.86 ± 0.08	0.96 ± 0.08	0.86 ± 0.04	0.92 ± 0.03
	RegNet	0.92 ± 0.05	0.91 ± 0.07	0.93 ± 0.08	0.84 ± 0.06	0.96 ± 0.03	0.85 ± 0.05	0.92 ± 0.03
	SwiftFormer	0.92 ± 0.06	0.89 ± 0.05	0.95 ± 0.04	0.84 ± 0.02	0.94 ± 0.08	0.85 ± 0.04	0.92 ± 0.05

**Table 6:** Quantitative evaluation of the proposed FDG-PET enhancement method in MCI classification task. Evaluated on Enhanced data. Values are reported as Mean ± Std.

Plane	Model	ACC	SEN	SPE	MCC	AUC	CK	F1
Sagittal	ViT	0.92 ± 0.04	0.92 ± 0.07	0.93 ± 0.06	0.85 ± 0.08	0.96 ± 0.03	0.84 ± 0.09	0.92 ± 0.06
	ConvNextV2	0.93 ± 0.04	0.88 ± 0.07	0.97 ± 0.04	0.86 ± 0.05	0.97 ± 0.03	0.86 ± 0.06	0.92 ± 0.02
	EfficientNetV2	0.92 ± 0.05	0.93 ± 0.12	0.91 ± 0.04	0.84 ± 0.08	0.97 ± 0.07	0.80 ± 0.04	0.92 ± 0.04
	GhostNetV3	0.90 ± 0.08	0.91 ± 0.10	0.90 ± 0.07	0.81 ± 0.12	0.95 ± 0.08	0.80 ± 0.07	0.90 ± 0.08
	MambaOut	0.92 ± 0.05	0.86 ± 0.07	0.98 ± 0.03	0.84 ± 0.06	0.96 ± 0.05	0.87 ± 0.05	0.91 ± 0.05
	MobileNetV4	0.93 ± 0.12	0.92 ± 0.08	0.94 ± 0.15	0.86 ± 0.06	0.97 ± 0.05	0.88 ± 0.11	0.93 ± 0.07
	RegNet	0.91 ± 0.07	0.90 ± 0.02	0.93 ± 0.04	0.83 ± 0.06	0.96 ± 0.09	0.87 ± 0.04	0.91 ± 0.06
	SwiftFormer	0.90 ± 0.07	0.88 ± 0.04	0.91 ± 0.07	0.80 ± 0.07	0.93 ± 0.11	0.79 ± 0.07	0.90 ± 0.08
Coronal	ViT	0.92 ± 0.07	0.90 ± 0.04	0.95 ± 0.03	0.85 ± 0.16	0.95 ± 0.09	0.87 ± 0.09	0.92 ± 0.06
	ConvNextV2	0.94 ± 0.03	0.89 ± 0.07	0.98 ± 0.04	0.88 ± 0.06	0.96 ± 0.03	0.85 ± 0.10	0.93 ± 0.08
	EfficientNetV2	0.93 ± 0.03	0.90 ± 0.06	0.96 ± 0.10	0.86 ± 0.09	0.96 ± 0.08	0.87 ± 0.04	0.93 ± 0.07
	GhostNetV3	0.91 ± 0.05	0.88 ± 0.12	0.94 ± 0.07	0.82 ± 0.06	0.96 ± 0.09	0.83 ± 0.03	0.91 ± 0.04
	MambaOut	0.92 ± 0.05	0.88 ± 0.08	0.96 ± 0.05	0.85 ± 0.04	0.96 ± 0.08	0.88 ± 0.08	0.92 ± 0.09
	MobileNetV4	0.93 ± 0.06	0.93 ± 0.06	0.94 ± 0.08	0.87 ± 0.11	0.97 ± 0.05	0.89 ± 0.06	0.93 ± 0.14
	RegNet	0.91 ± 0.08	0.89 ± 0.07	0.93 ± 0.05	0.82 ± 0.10	0.97 ± 0.09	0.83 ± 0.08	0.91 ± 0.08
	SwiftFormer	0.92 ± 0.10	0.89 ± 0.05	0.95 ± 0.06	0.84 ± 0.07	0.96 ± 0.04	0.83 ± 0.04	0.92 ± 0.05
Axial	ViT	0.92 ± 0.05	0.91 ± 0.08	0.94 ± 0.05	0.84 ± 0.04	0.96 ± 0.07	0.85 ± 0.03	0.92 ± 0.09
	ConvNextV2	0.93 ± 0.06	0.90 ± 0.07	0.96 ± 0.03	0.86 ± 0.08	0.97 ± 0.09	0.86 ± 0.06	0.93 ± 0.08
	EfficientNetV2	0.94 ± 0.07	0.92 ± 0.06	0.96 ± 0.05	0.89 ± 0.09	0.98 ± 0.06	0.86 ± 0.05	0.94 ± 0.08
	GhostNetV3	0.92 ± 0.11	0.92 ± 0.05	0.92 ± 0.05	0.84 ± 0.11	0.97 ± 0.07	0.84 ± 0.05	0.92 ± 0.04
	MambaOut	0.95 ± 0.07	0.94 ± 0.08	0.96 ± 0.04	0.91 ± 0.08	0.98 ± 0.09	0.94 ± 0.03	0.95 ± 0.01
	MobileNetV4	0.96 ± 0.05	0.95 ± 0.06	0.96 ± 0.15	0.92 ± 0.09	0.99 ± 0.07	0.90 ± 0.10	0.96 ± 0.08
	RegNet	0.94 ± 0.06	0.93 ± 0.09	0.95 ± 0.10	0.88 ± 0.08	0.99 ± 0.03	0.89 ± 0.05	0.94 ± 0.04
	SwiftFormer	0.93 ± 0.06	0.91 ± 0.07	0.94 ± 0.05	0.85 ± 0.08	0.97 ± 0.09	0.89 ± 0.07	0.93 ± 0.06

When comparing the two tables, a consistent pattern emerges: Hybrid enhancement narrows the performance gap between architectures and stabilizes metric variability across planes. In raw data, mean accuracy ranges span up to 7 percentage points for some planes, whereas enhanced data reduces these spreads, suggesting a more uniform and separable feature space. Sensitivity gains are particularly relevant from a clinical perspective, with MambaOut (Axial) improving from  $0.89 \pm 0.07$  to  $0.94 \pm 0.08$ , indicating better detection of the subtle hypometabolism that characterizes early-stage MCI. AUC values also show systematic gains, frequently reaching 0.97–0.99 for axial models, reflecting stronger class separation independent of threshold selection. Taken together, these results reinforce the central conclusion of this work: the limiting factor in automated MCI classification is not the model architecture but the quality of the input signal.

By enhancing structural boundaries while preserving noise statistics, the Hybrid method provides a more discriminative and radiomic-stable representation, thereby improving downstream diagnostic performance.

## 5 Discussion

The results of this study show that the proposed Hybrid Laplacian-DoG framework mitigates the longstanding trade-off between contrast enhancement and noise amplification in FDG-PET neuroimaging. By combining a structure-preserving Laplacian base with a gradient-gated DoG detail injector, the method produces a balanced enhancement profile ( $NG \approx 1.01$ ,  $\Delta\text{CNR} > 0$ ) that directly improves downstream MCI classification performance.

A key finding is that widely used enhancement techniques are poorly suited to the statistical properties of PET data. As illustrated in Fig. 4, Bilateral filtering and CLAHE implicitly assume that local smoothness or histogram redistribution will reduce noise. However, cortical FDG uptake is inherently textural due to underlying neurophysiology and Poisson acquisition noise. The enforcement of smoothness (Bilateral) or alteration of the global intensity distribution (CLAHE) disrupts these subtle patterns, leading to elevated Noise Gain values ( $>1.25$ ) and degraded structural fidelity.

The Hybrid framework performs well precisely because it aligns with the modality's physics. The Laplacian stage enhances the underlying anatomical structure without modifying the noise floor. Subsequent DoG injection is regulated by the gradient-based gating term  $W_{grad}$ , which ensures that sharpening is applied only at locations with genuine radiotracer gradients. This behavior effectively suppresses Poisson noise amplification in homogeneous white matter and ventricular regions while restoring structure lost to partial-volume effects.

From a clinical perspective, an enhancement method must not introduce artifacts that resemble pathology. Noise residual analysis (Fig. 8) was essential to evaluate this aspect. The Hybrid method produced sparse, anatomically aligned residuals, whereas Unsharp Masking exhibited diffuse speckling and Bilateral filtering showed clustered intensity shifts, which indicates artificial texture generation. The structural coherence of the Hybrid residuals underscores its safety and reliability for diagnostic workflows.

The favorable balance between noise and structure likely underlies the performance gains observed for MobileNetV4 and MambaOut. Deep learning classifiers are sensitive to high-frequency corruption. The Hybrid method provides cleaner gradients and more coherent boundaries, yielding a more separable feature space. The reduction in misclassification error on the axial plane illustrates the practical diagnostic benefit of this approach.

The consistent superiority of the axial plane is also notable. The axial slices align with the native acquisition geometry of the PET scanners, minimizing interpolation artifacts. In addition, early hypometabolism associated with Alzheimer's disease and MCI is most prominent along the cortical ribbon in axial views. The Hybrid method is particularly effective in enhancing these subtle patterns, contributing to the observed 96% mean accuracy in axial classification.

It is important to note that EPI quantifies the spatial correspondence of edges between the raw and enhanced images rather than the magnitude of contrast amplification. A high EPI value indicates that the enhancement preserves the location and orientation of the anatomical boundaries; it does not imply that no enhancement has occurred. Instead, it confirms that the method does not introduce artificial edges or alter the geometric structure of the tracer distribution. The contrast enhancement primarily modifies the strength of the gradients, while the EPI evaluates the pattern of those gradients. Consequently, methods such as Laplacian filtering can achieve EPI scores near 0.99 while still producing meaningful local contrast increases without distorting anatomical geometry.

Although a dedicated ablation table is not provided, the experimental design serves as an implicit ablation study, as each module of the framework is evaluated independently. The results for the Laplacian-only variant, the DoG-only variant, and the full Hybrid method isolate the contributions of edge-preserving contrast modulation, high-frequency detail extraction, and their combined effect. The behavior of the gradient-gating term is evident from the contrast–noise relationship: the DoG baseline increases noise and reduces edge fidelity, whereas the Hybrid method restores contrast while keeping Noise Gain near unity. Similarly, intensity-weighting suppresses enhancement in low-uptake regions, preventing noise amplification in homogeneous tissue.

Despite these promising results, several limitations remain. While we did not perform an exhaustive sensitivity analysis over the full combinatorial space of all parameters, we conducted a targeted sensitivity study on the primary modulation parameters that directly control enhancement strength, namely the Laplacian gain  $k$ , the DoG mixing factor  $\alpha$ , and the gradient gate threshold  $\tau$ . These parameters were varied within physiologically and physically plausible ranges, informed by PET spatial resolution, cortical thickness, and normalized gray-matter intensity distributions. Across all tested configurations, the Hybrid method exhibited stable behavior, with consistently positive contrast gains ( $\Delta\text{CNR}$ ), high edge preservation (EPI close to unity), and noise gain remaining near one. This indicates that the proposed framework is robust to moderate parameter variation and does not rely on fine-tuned operating points to achieve its reported performance. Parameters related to normalization (e.g., percentile anchors) were not included in sensitivity tuning, as they are deterministic functions of the input distribution rather than free hyperparameters. A more exhaustive exploration of the parameter space, potentially incorporating cross-validation, scanner-specific stratification, or Bayesian optimization, could further quantify robustness and identify acquisition-dependent operating points. Such an analysis, particularly in large multi-center or low-dose PET settings, is left for future work.

The clinical relevance of the proposed enhancement was evaluated using 2D mid-slices extracted from enhanced FDG-PET volumes, while the enhancement itself operates fully in 3D. This evaluation protocol was intentionally adopted to isolate the effect of volumetric enhancement on diagnostic signal quality, independently of confounding factors introduced by 3D classifier capacity and architectural variability. Importantly, 2D slice-based classification remains a common and accepted practice in FDG-PET studies of Alzheimer's disease and MCI, with numerous prior works reporting clinically meaningful results under the same setting [39–43]. All classifier inputs in this study are derived from fully enhanced 3D volumes. Therefore, the observed performance gains directly reflect improvements in the underlying volumetric signal rather than slice-wise processing artifacts. While end-to-end 3D classification may further exploit spatial context, it represents a complementary research direction and is left for future work.

A further limitation relates to the interpretation of the NG metric. NG provides an image-domain estimate of high-frequency variance within regions that appear homogeneous in the raw volume and is therefore intended for relative comparison of enhancement methods under identical acquisition settings, rather than for absolute characterization of scanner noise. The percentile-based definition of “flat tissue” (voxels below the 30th percentile of the raw gradient magnitude) serves as a practical surrogate commonly used in PET image-processing studies when physical noise measurements are unavailable. In FDG-PET, this threshold reliably identifies the ventricular CSF and the deep white matter, which are physiologically low-variance regions. Nevertheless, NG remains an indirect proxy for noise behavior, and future work should incorporate phantom experiments or scanner-native noise descriptors to provide a more direct and physics-grounded assessment of noise amplification.

Regarding quantitative PET measures such as SUV, we emphasize that the proposed framework is designed as a structure-preserving enhancement rather than a quantitative correction method. Global intensity scaling is explicitly preserved by median normalization, and NG values close to unity indicate

that enhancement does not alter the noise floor in homogeneous tissue. However, because the method locally modulates gradient strength to restore anatomical boundaries, small voxel-level changes in SUV values near edges may occur. Importantly, the present work does not claim strict voxel-wise SUV invariance, particularly in regions affected by partial-volume effects. Instead, the method aims to improve anatomical fidelity and discriminative signal quality while maintaining global quantitative consistency. A dedicated statistical evaluation of SUV stability, including region-wise analysis and phantom validation, is an important direction for future work.

It is also important to clarify that although several components of the framework rely on data-derived percentiles (e.g., for edge normalization or intensity weighting), these operations do not introduce stochasticity. All percentile computations are deterministic functions of the input volume and produce identical results for identical inputs. Their role is to adapt the enhancement strength to the subject's intensity distribution, analogously to widely used histogram-based normalization procedures. Thus, the framework remains fully deterministic in practice, despite employing data-dependent weighting terms.

A further limitation of this study is the absence of reconstruction-level methods in the comparison. Techniques such as point-spread-function (PSF) modeling, advanced OSEM variants with regularization, highly constrained back-projection (HYPR) based denoising, and proprietary deep-learning reconstructions (e.g., Siemens AI.PET) operate directly on sinogram data or incorporate scanner-specific system matrices. These approaches differ fundamentally in scope and are not directly comparable to our framework, which is a post-processing enhancement method applied to fully reconstructed standard-dose FDG-PET volumes. As such, the present evaluation focuses exclusively on image-domain techniques that can be deployed independently of scanner hardware or vendor-specific reconstruction pipelines.

No formal qualitative reader study involving nuclear medicine physicians was conducted in this work. The evaluation was based on objective, reproducible quantitative metrics and standardized visual comparisons to avoid subjective bias. Expert reader assessment represents an important direction for future validation.

Furthermore, although ADNI provides high-quality standardized data, real-world clinical scans vary widely in dose, reconstruction kernels, and acquisition protocols. Evaluating robustness across multi-site, heterogeneous datasets will be an essential next step.

## 6 Conclusions

This work introduced a 3D enhancement framework designed specifically for the characteristics of FDG-PET neuroimaging. By combining the structural stability of a Laplacian operator with the contrast enhancement provided by a gradient-gated Difference-of-Gaussians, the proposed Hybrid method effectively resolves the conventional trade-off between image sharpness and noise amplification.

The evaluation confirms that the method offers a practical and reliable enhancement strategy. Quantitatively, it is the only approach tested that achieves significant contrast improvement ( $\Delta\text{CNR} > 0$ ) while maintaining a near-unity Noise Gain ( $NG \approx 1.01$ ), indicating that structural contrast increases without altering the underlying noise distribution. Qualitatively, the method restores fine anatomical detail that is often suppressed by partial volume effects in raw PET data.

The observed improvements in the detection of mild cognitive impairment further highlight its potential in clinical relevance. Without modifying network architectures, hybrid enhancement reduced the mean misclassification rate from approximately 7% to 4%, corresponding to a roughly 40% reduction in errors. These findings suggest that a substantial part of the challenge in automated MCI detection stems from limitations in the input signal rather than the model's capacity. Enhancing PET data in a

structurally faithful and noise-stable manner may therefore be a critical step toward more accurate and reliable computer-aided diagnosis.

**Acknowledgement:** None.

**Funding Statement:** The authors received no specific funding for this study.

**Author Contributions:** Conceptualization, Ovidijus Grigas; Data curation, Rytis Maskeliūnas; Formal analysis, Ovidijus Grigas and Rytis Maskeliūnas; Funding acquisition, Rytis Maskeliūnas; Investigation, Ovidijus Grigas and Rytis Maskeliūnas; Methodology, Ovidijus Grigas; Project administration, Rytis Maskeliūnas; Resources, Ovidijus Grigas; Software, Ovidijus Grigas; Supervision, Rytis Maskeliūnas; Validation, Ovidijus Grigas; Visualization, Ovidijus Grigas; Writing—original draft, Ovidijus Grigas; Writing—review & editing, Ovidijus Grigas and Rytis Maskeliūnas. All authors reviewed and approved the final version of the manuscript.

**Availability of Data and Materials:** Dataset Alzheimer's Disease Neuroimaging Initiative (ADNI) used in this study can be accessed upon request through IDA system <https://ida.loni.usc.edu> (accessed on 22 November 2025).

**Ethics Approval:** Not applicable.

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

## References

1. Anderson ND. State of the science on mild cognitive impairment (MCI). *CNS Spectr*. 2019;24(1):78–87. doi:10.1017/s1092852918001347.
2. Pérot JB, Niewiadomska-Cimicka A, Brouillet E, Trottier Y, Flament J. Longitudinal MRI and <sup>1</sup>H-MRS study of SCA7 mouse forebrain reveals progressive multiregional atrophy and early brain metabolite changes indicating early neuronal and glial dysfunction. *PLoS One*. 2024;19(1):e0296790. doi:10.1371/journal.pone.0296790.
3. Mosconi L. Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease: FDG-PET studies in MCI and AD. *Eur J Nucl Med Mol Imaging*. 2005;32(4):486–510. doi:10.1007/s00259-005-1762-7.
4. Ibaraki M, Matsubara K, Shinohara Y, Shidahara M, Sato K, Yamamoto H, et al. Brain partial volume correction with point spreading function reconstruction in high-resolution digital PET: comparison with an MR-based method in FDG imaging. *Ann Nucl Med*. 2022;36(8):717–27. doi:10.1007/s12149-022-01753-5.
5. Vigneron V, Kodewitz A, Tome AM, Lelandais S, Lang E. Alzheimer's disease brain areas: the machine learning support for blind localization. *Curr Alzheimer Res*. 2016;13(5):498–508. doi:10.2174/1567205013666160314144822.
6. Rahmim A, Tang J. Noise propagation in resolution modeled PET imaging and its impact on detectability. *Phys Med Biol*. 2013;58(19):6945. doi:10.1088/0031-9155/58/19/6945.
7. Harada K, Ohashi Y, Chiba A, Numasawa K, Imai T, Hayasaka S, et al. Development of new digital phantom creation tool for evaluation of low-contrast detectability using iterative reconstruction. *Nippon Hoshasen Gijutsu Gakkai Zasshi*. 2018;74(8):769–78. doi:10.6009/jjrt.2018\_JSRT\_74.8.769.
8. Mittal D, Kumar V, Saxena SC, Khandelwal N, Kalra N. Enhancement of the ultrasound images by modified anisotropic diffusion method. *Med Biol Eng Comput*. 2010;48(12):1281–91. doi:10.1007/s11517-010-0650-x.
9. Vigneshwaran B, Maheswari RV, Subburaj P. An improved threshold estimation technique for partial discharge signal denoising using Wavelet Transform. In: *Proceedings of the 2013 International Conference on Circuits, Power and Computing Technologies (ICCPCT)*; 2013 Mar 20–21; Nagercoil, India. p. 300–5. doi:10.1109/iccpct.2013.6528823.
10. Kinahan PE, Fletcher JW. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. *Semin Ultrasound CT MRI*. 2010;31(6):496–505. doi:10.1053/j.sult.2010.10.001.
11. Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, et al. Photo-realistic single image super-resolution using a generative adversarial network. *arXiv:1609.04802v5*. 2016. doi:10.48550/arXiv.1609.0480202.

12. Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv:1703.10593v7*. 2017. doi:10.48550/arXiv.1703.10593.
13. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *arXiv:1505.04597*. 2015. doi: 10.48550/arXiv.1505.04597.
14. Lei Y, Harms J, Wang T, Liu Y, Shu HK, Jani AB, et al. MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks. *Med Phys*. 2019;46(8):3565–81. doi:10.1002/mp.13617.
15. Tang C, Li J, Wang L, Li Z, Jiang L, Cai A, et al. Unpaired low-dose CT denoising network based on cycle-consistent generative adversarial network with prior image information. *Comput Math Meth Med*. 2019;2019:8639825. doi:10.1155/2019/8639825.
16. Di Feola F, Pompilio L, Assolito C, Guarrasi V, Soda P. Texture-aware StarGAN for CT data harmonization. In: *Proceedings of the 2025 International Joint Conference on Neural Networks (IJCNN)*; 2025 Jun 30–Jul 5; Rome, Italy. p. 1–8. doi:10.1109/ijcnn64981.2025.11228038.
17. Marey A, Arjmand P, Alerab ADS, Eslami MJ, Saad AM, Sanchez N, et al. Explainability, transparency and black box challenges of AI in radiology: impact on patient care in cardiovascular radiology. *Egypt J Radiol Nucl Med*. 2024;55(1):183. doi:10.1186/s43055-024-01356-2.
18. Flaus A, Deddah T, Reilhac A, De Leiris N, Janier M, Merida I, et al. PET image enhancement using artificial intelligence for better characterization of epilepsy lesions. *Front Med*. 2022;9:1042706. doi:10.3389/fmed.2022.1042706.
19. Chaudhari AS, Mitra E, Davidzon GA, Gulaka P, Gandhi H, Brown A, et al. Low-count whole-body PET with deep learning in a multicenter and externally validated study. *npj Digit Med*. 2021;4:127. doi:10.1038/s41746-021-00497-2.
20. Song TA, Chowdhury SR, Yang F, Dutta J. PET image super-resolution using generative adversarial networks. *Neural Netw*. 2020;125:83–91. doi:10.1016/j.neunet.2020.01.029.
21. Hashimoto F, Ohba H, Ote K, Teramoto A, Tsukada H. Dynamic PET image denoising using deep convolutional neural networks without prior training datasets. *IEEE Access*. 2019;7:96594–603. doi:10.1109/access.2019.2929230.
22. Jiang C, Liu M, Sun K, Shen D. End-to-end triple-domain pet enhancement: a hybrid denoising-and-reconstruction framework for reconstructing standard-dose PET images from low-dose PET sinograms. *arXiv:2412.03617*. 2024. doi:10.48550/arXiv.2412.03617.
23. Xue H, Zhang Q, Zou S, Zhang W, Zhou C, Tie C, et al. LCPR-Net: low-count PET image reconstruction using the domain transform and cycle-consistent generative adversarial networks. *Quant Imaging Med Surg*. 2020;11(2):749–62. doi:10.21037/qims-20-66.
24. Yoshimura T, Hasegawa A, Kogame S, Magota K, Kimura R, Watanabe S, et al. Medical radiation exposure reduction in PET via super-resolution deep learning model. *Diagnostics*. 2022;12(4):872. doi:10.3390/diagnostics12040872.
25. Chen Z, Yang H, Qi M, Chen W, Liu F, Song S, et al. Enhancing 18F-FDG PET image quality and lesion diagnostic performance across different body mass index using the deep progressive learning reconstruction algorithm. *Cancer Imag*. 2025;25(1):58. doi:10.1186/s40644-025-00877-x.
26. Hoopes A, Mora JS, Dalca AV, Fischl B, Hoffmann M. SynthStrip: skull-stripping for any brain image. *NeuroImage*. 2022;260(1):119474. doi:10.1016/j.neuroimage.2022.119474.
27. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *NeuroImage*. 2012;62(2):782–90. doi:10.1016/j.neuroimage.2011.09.015.
28. Vinoothna B, Rajendiran N. Design and development of contrast-limited adaptive histogram equalization technique for enhancing pet images by improving joint entropy, UIQI parameters in comparison with median filtering. *AIP Conf Proc*. 2024;2816:090014. doi:10.1063/5.0185943.
29. Zuiderveld K. Contrast limited adaptive histogram equalization. San Diego, CA, USA: Academic Press Professional, Inc.; 1994. p. 474–85.
30. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An image is worth 16 × 16 words: Transformers for image recognition at scale. *arXiv:2010.11929*. 2020. doi:10.48550/arXiv.2010.11929.

31. Shaker A, Maaz M, Rasheed H, Khan S, Yang MH, Khan FS. SwiftFormer: efficient additive attention for transformer-based real-time mobile vision applications. *arXiv:2303.15446*. 2023. doi:10.48550/arXiv.2303.15446.
32. Woo S, Debnath S, Hu R, Chen X, Liu Z, Kweon IS, et al. ConvNeXt V2: co-designing and scaling ConvNets with masked autoencoders. *arXiv:2301.00808*. 2023. doi:10.48550/arXiv.2301.00808.
33. Tan M, Le QV. EfficientNetV2: smaller models and faster training. *arXiv:2104.00298*. 2021. doi:10.48550/arXiv.2104.00298.
34. Liu Z, Hao Z, Han K, Tang Y, Wang Y. GhostNetV3: exploring the training strategies for compact models. *arXiv:2404.11202*. 2024. doi:10.48550/arXiv.2404.11202.
35. Qin D, Lechner C, Delakis M, Fornoni M, Luo S, Yang F, et al. MobileNetV4—universal models for the mobile ecosystem. *arXiv:2404.10518*. 2024. doi:10.48550/arXiv.2404.10518.
36. Xu J, Pan Y, Pan X, Hoi S, Yi Z, Xu Z. RegNet: self-regulated network for image classification. *arXiv:2101.00590*. 2021. doi:10.48550/arXiv.2101.00590.
37. Yu W, Wang X. MambaOut: do we really need mamba for vision? *arXiv:2405.07992*. 2024. doi:10.48550/arXiv.2405.07992.
38. Veraart J, Novikov DS, Christiaens D, Ades-aron B, Sijbers J, Fieremans E. Denoising of diffusion MRI using random matrix theory. *NeuroImage*. 2016;142(4):394–406. doi:10.1016/j.neuroimage.2016.08.016.
39. Liu M, Cheng D, Yan W, Initiative ADN. Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images. *Front Neuroinform*. 2018;12:35. doi:10.3389/fninf.2018.00035.
40. Zhang F, Li Z, Zhang B, Du H, Wang B, Zhang X. Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease. *Neurocomputing*. 2019;361(5):185–95. doi:10.1016/j.neucom.2019.04.093.
41. Kim HW, Lee HE, Oh K, Lee S, Yun M, Yoo SK. Multi-slice representational learning of convolutional neural network for Alzheimer's disease classification using positron emission tomography. *Biomed Eng Online*. 2020;19(1):70. doi:10.1186/s12938-020-00813-z.
42. Pan X, Phan TL, Adel M, Fossati C, Gaidon T, Wojak J, et al. Multi-view separable pyramid network for AD prediction at MCI stage by 18F-FDG brain PET imaging. *IEEE Trans Med Imaging*. 2021;40(1):81–92. doi:10.1109/tmi.2020.3022591.
43. Rehman A, Yi MK, Majeed A, Hwang SO. Early diagnosis of Alzheimer's disease using 18F-FDG PET with soften latent representation. *IEEE Access*. 2024;12(11):87923–33. doi:10.1109/access.2024.3418508.