



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Unsupervised Brain Anomaly Detection as a Bayesian Inverse Problem with Diffusion Prior

Hugues Roy¹[0009-0006-3503-6849], Reuben Dorent^{1*}[0000-0002-7530-0644], and
Ninon Burgos^{1*}[0000-0002-4668-2006]

Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inria,
Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013, Paris, France

*Contributed equally
huguesroy1277@gmail.com

Abstract. Unsupervised anomaly detection (UAD) aims to localize abnormal regions in medical scans without pixel-level annotations. A typical strategy seeks to reconstruct a pseudo-healthy image that preserves subject-specific anatomy. Recently, diffusion models have been proposed to perform UAD. However, these methods rely on heuristic noise schedules or synthetic corruptions to balance subject-specificity and anomaly removal. In this work, we propose an alternative formulation of UAD as a Bayesian inverse problem under a diffusion prior. First, we introduce a latent spatial anomaly mask that models pixel-wise consistency between a test image and its latent corresponding pseudo-healthy image. Then, we propose an approximation of the unknown generation process that links healthy anatomy, anomalies, and the observed image, enabling a well-defined likelihood within the Bayesian framework. Building on recent advances in diffusion-based inverse problem methods, we jointly infer the pseudo-healthy image and the anomaly mask via annealed posterior sampling. We evaluate our approach on FDG PET (ADNI) and FLAIR MRI (BraTS 2021), demonstrating improved anomaly localization performance compared to other diffusion-based approaches and validating the contribution of our introduced model. Our code is available at https://github.com/HuguesRoy/UAD_DAPS.

Keywords: Unsupervised anomaly detection · Inverse problem

1 Introduction

Unsupervised anomaly detection (UAD) in medical imaging aims to identify abnormal regions without requiring assumptions about anomaly characteristics or pixel-level annotations. A common strategy is to reconstruct a pseudo-healthy version of an abnormal image to detect anomalies.

Such reconstruction must satisfy two competing objectives: (i) it should be plausible under the healthy data distribution, and (ii) it should preserve subject-specific anatomical details.

Diffusion models trained on healthy data have recently been used for UAD due to their strong generative capabilities. However, to balance these competing

objectives, these methods rely on heuristic choices, such as denoising noise levels [1,3,4,23], corruption strategies [4,22,24,18], or synthetic anomalies [11,13,2].

In parallel, diffusion models recently achieved strong performance in Bayesian inverse problems by leveraging learned data priors [19,25]. These approaches progressively estimate a latent clean signal from noisy observations through iterative denoising updates. However, they rely on a known forward model linking the clean signal to the noisy measurement, which is unavailable in UAD.

In this work, we bridge these two lines of research by formulating unsupervised anomaly detection as a Bayesian inverse problem in which the forward model is unknown and must be approximated. Our contributions are fourfold: (1) we formulate diffusion-based UAD as a Bayesian inverse problem with a latent anomaly mask; (2) we introduce an approximation of the unknown forward model; (3) we propose a spatially regularized mask prior enabling stable joint inference via intermediate posterior sampling [25]; and (4) we conduct extensive experiments demonstrating the effectiveness of our approach which outperforms existing diffusion-based methods on both PET and MRI benchmarks.

2 Inverse problem using diffusion models

Inverse problems aim to reconstruct an unknown signal $\mathbf{x}_0$ from a partial, noisy measurement $\mathbf{y}$ by inverting a known *forward model* that describes the measurement process. In a Bayesian setting, reconstruction is formulated as inference under the posterior distribution $p(\mathbf{x}_0 | \mathbf{y}) \propto p(\mathbf{y} | \mathbf{x}_0)p(\mathbf{x}_0)$, where the likelihood $p(\mathbf{y} | \mathbf{x}_0)$ is defined by the forward model and $p(\mathbf{x}_0)$ is the prior distribution. A key challenge of inverse problems is to approximate this unknown prior.

Recently, diffusion models have shown strong performance to approximate priors [5,6,7,12,19] in inverse problems. A noisy signal $\mathbf{x}_t$ is iteratively denoised using both a prior diffusion model trained on clean signals and the likelihood $p(\mathbf{y} | \mathbf{x}_t)$, leading to a denoising process conditioned on the measurement $\mathbf{y}$. However, early denoising errors have been found to be difficult to correct [25].

To address this limitation, Zhang et al. [25] proposed the Decoupled Annealed Posterior Sampling (DAPS) strategy. Rather than strictly following the diffusion trajectory, DAPS introduces a sequence of intermediate posteriors $p_k(\mathbf{x}_0 | \mathbf{y}, \mathbf{x}_{t_k}) \propto p(\mathbf{y} | \mathbf{x}_0)p_k(\mathbf{x}_0 | \mathbf{x}_{t_k})$ which gradually converge to the true posterior $p(\mathbf{x}_0 | \mathbf{y})$ and enable exploring a larger solution space. As in [6,19], they approximate the conditional prior $p_k(\mathbf{x}_0 | \mathbf{x}_{t_k})$ using a Gaussian distribution,

$$p_k(\mathbf{x}_0 | \mathbf{x}_{t_k}) \approx \mathcal{N}(\mathbf{x}_0; \hat{\mathbf{x}}_0(\mathbf{x}_{t_k}, t_k), r_k^2 \mathbf{I}), \quad (1)$$

where $\hat{\mathbf{x}}_0$ is sampled from the reverse process of the diffusion model. The variance r_k^2 weights the contribution of the conditional prior in the intermediate posterior.

Instead of directly sampling $\mathbf{x}_{t_{k+1}}$ from $\mathbf{x}_{t_k}$, an intermediate signal $\mathbf{x}_0^{(k)}$ is sampled using the intermediate posterior $p_k(\mathbf{x}_0 | \mathbf{y}, \mathbf{x}_{t_k})$ and re-noised to obtain a sample $\mathbf{x}_{t_{k+1}}$. Thanks to this decoupled sampling strategy and intermediate posterior parameterization, larger changes between iterations are obtained, preventing early denoising errors from propagating through the sampling process.

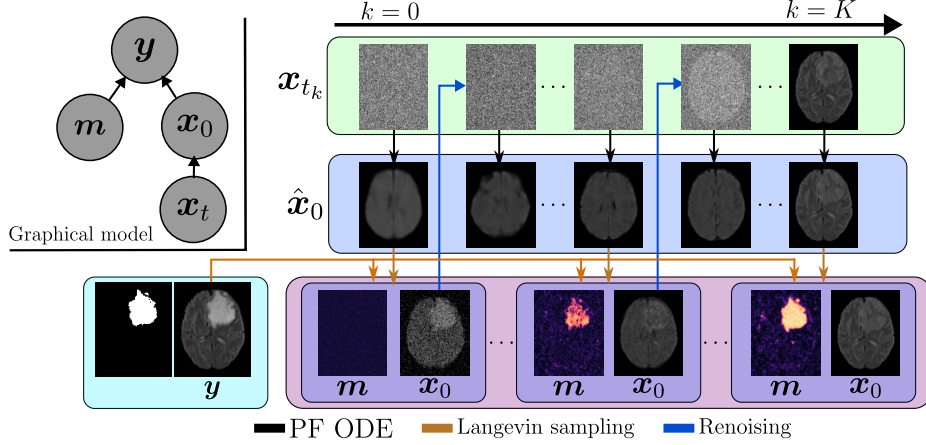


Fig. 1. Top left: Probabilistic graphical model of the proposed method. **Right:** Overview of the inference process. As k increases, the posterior distribution progressively concentrates, enabling the sampling of the final estimates x_0 and m . The variable x_{t_k} denotes the noisy version of x_0 at step k , which allows a decoupling between the previously sampled noisy variable and the newly generated one.

A key limitation of DAPS and other inverse problem approaches is their reliance on a known and available forward model. This assumption typically does not hold in unsupervised anomaly detection for medical imaging, where no forward model exists to sample an abnormal image y from a healthy image x_0 .

3 Method

In this work, we propose a new diffusion-based approach to unsupervised anomaly detection. Specifically, we formulate anomaly detection as a Bayesian inverse problem and approximate the unknown forward model by introducing a latent variable m that represents a probabilistic anomaly mask. Under this formulation, the objective is to jointly infer the healthy image x_0 and the mask m from the observed image y . In the following sections, we define the probabilistic model, including the likelihood, the priors, and the proposed inference strategy.

Generative model To formalize anomaly detection as a Bayesian inverse problem, we assume that the observed image y is generated from a latent healthy image x_0 and a latent anomaly mask m through the likelihood $p(y | x_0, m)$. Specifically, the observation y is expected to coincide with x_0 only in regions labeled as normal by the mask m . As shown in the graphical model in Fig. 1, assuming prior independence between x_0 and m , the joint distribution factorizes as $p(x_0, m, y) = p(x_0)p(m)p(y | x_0, m)$, where $p(x_0)$ is the prior over healthy images and $p(m)$ is a spatial prior over anomaly masks. As in Bayesian inverse problems, the objective is to sample the healthy image x_0 and anomaly mask m .

from the posterior $p(\mathbf{x}_0, \mathbf{m} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{x}_0, \mathbf{m})p(\mathbf{x}_0)p(\mathbf{m})$. To approximate this posterior, we adopt a similar formulation as in DAPS [25]. Introducing the noisy diffusion states $\mathbf{x}_{t_k}$ to approximate the image prior $p(\mathbf{x}_0)$, we define a sequence of intermediate posteriors

$$p_k(\mathbf{x}_0, \mathbf{m} \mid \mathbf{y}, \mathbf{x}_{t_k}) \propto p_k(\mathbf{y} \mid \mathbf{x}_0, \mathbf{m})p_k(\mathbf{x}_0 \mid \mathbf{x}_{t_k})p(\mathbf{m}), \quad (2)$$

where $p_k(\mathbf{x}_0 \mid \mathbf{x}_{t_k})$ is the conditional diffusion prior defined in Eq. (1).

Approximate the forward model Unlike classical inverse problems, the forward model relating a healthy image $\mathbf{x}_0$ to an abnormal scan $\mathbf{y}$ is unknown. In this section, we propose an approximation of the forward model $p(\mathbf{y} \mid \mathbf{x}_0, \mathbf{m})$ based on the introduced anomaly mask $\mathbf{m}$. The probabilistic anomaly mask $\mathbf{m}$ encodes pixel-level abnormality: values close to 0 indicate anomalous measurements, whereas values close to 1 enforce consistency between $\mathbf{y}$ and $\mathbf{x}_0$. Accordingly, we model the likelihood as:

$$p_k(\mathbf{y} \mid \mathbf{x}_0, \mathbf{m}) = \mathcal{N}(\mathbf{y}; \mathbf{x}_0, r_k^2 \text{diag}(\mathbf{m})^{-2}) , \quad (3)$$

with variance r_k^2 , matched to the conditional prior $p_k(\mathbf{x}_0 \mid \mathbf{x}_{t_k})$ variance. This ensures that, at intermediate step k , in normal regions, the variability of the diffusion-based prediction is consistent with the variability observed in $\mathbf{y}$. Deviations exceeding this expected variability cannot be explained by the prior and are therefore attributed to abnormal regions.

Spatial prior on the mask Since the probabilistic mask $\mathbf{m}$ is introduced as a latent variable, we must specify a prior distribution $p(\mathbf{m})$. To enforce values in $(0, 1)$, we parametrize it as $\mathbf{m} = s(\mathbf{a})$, where $\mathbf{a}$ are unconstrained logits and s denotes the sigmoid function. We then define a prior $p(\mathbf{a})$, that (i) defines the initial mask value, (ii) controls the magnitude of mask updates during inference, and (iii) enforces spatial consistency. To this end, we assume a Gaussian Markov random field prior:

$$p(\mathbf{a}) = \mathcal{N}(\mathbf{a}; \mu \mathbf{1}, \mathbf{Q}^{-1}) , \quad (4)$$

with a precision matrix $\mathbf{Q} = \lambda_0 \mathbf{I} + \lambda_c \mathbf{L}$, where $\lambda_0, \lambda_c > 0$ and $\mathbf{L}$ is the graph Laplacian of the image grid. The quadratic form $\mathbf{a}^\top \mathbf{L} \mathbf{a}$ penalizes differences between neighboring logits, promoting spatial consistency. The parameter μ sets the initial value of the mask, λ_0 controls the global shrinkage of logits toward μ , and λ_c governs the spatial consistency strength.

Energy formulation of the intermediate posterior Finally, by combining the conditional diffusion prior defined in Eq. (1), the approximate likelihood in Eq. (3) and the logit mask prior in Eq. (4), the intermediate posteriors in

Eq. (2) can be formulated as a Gibbs distribution, i.e. $p_k(\mathbf{x}_0, \mathbf{a} \mid \mathbf{y}, \mathbf{x}_{t_k}) \propto \exp(-U_k(\mathbf{x}_0, \mathbf{a}))$, where $U_k(\mathbf{x}_0, \mathbf{a})$ is defined as:

$$\begin{aligned}
 U_k(\mathbf{x}_0, \mathbf{a}) = & \underbrace{\frac{1}{2r_k^2} \|\mathbf{x}_0 - \hat{\mathbf{x}}_0(\mathbf{x}_{t_k}, t_k)\|^2}_{\text{Conditional prior consistency}} + \underbrace{\frac{1}{2r_k^2} \|s(\mathbf{a}) \odot (\mathbf{y} - \mathbf{x}_0)\|^2}_{\text{Mask-weighted measurement consistency}} \\
 & - \underbrace{\sum_i \log s(a_i) + \frac{\lambda_0}{2} \|\mathbf{a} - \mu \mathbf{1}\|^2}_{\text{Mask value regularization}} + \underbrace{\frac{\lambda_c}{2} \sum_{(i,j) \in E} (a_i - a_j)^2}_{\text{Spatial mask consistency}}. \quad (5)
 \end{aligned}$$

Inference by Langevin dynamics In order to sample from the intermediate posteriors, we use unadjusted Langevin dynamics [21] targeting the energy U_k at each intermediate step k . We denote j the Langevin steps and $(\mathbf{x}_0^{(j)}, \mathbf{a}^{(j)})$ the samples at step j . At each step k , the reconstruction $\mathbf{x}_0$ is initialized using the mean of the conditional prior sample and the observation. The mask logits are initialized once at $k = 0$ as $\mathbf{a}^0 = \mu \mathbf{1}$ and are propagated across intermediate steps without re-initialization, enabling progressive refinement of the anomaly mask as the noise level decreases.

$$\mathbf{x}_0^{(j+1)} = \mathbf{x}_0^{(j)} - \eta_x \nabla_{\mathbf{x}_0} U_k(\mathbf{x}_0^{(j)}, \mathbf{a}^{(j)}) + \sqrt{2\eta_x} \boldsymbol{\xi}_x^{(j)}, \quad (6)$$

$$\mathbf{a}^{(j+1)} = \mathbf{a}^{(j)} - \eta_a \nabla_{\mathbf{a}} U_k(\mathbf{x}_0^{(j)}, \mathbf{a}^{(j)}) + \sqrt{2\eta_a} \boldsymbol{\xi}_a^{(j)}, \quad (7)$$

where $\boldsymbol{\xi}_x^{(j)}$ and $\boldsymbol{\xi}_a^{(j)}$ are standard Gaussian noise terms. We use only a few iterations per intermediate step k . The step sizes $\eta_x, \eta_a > 0$ are automatically adjusted based on gradient magnitudes $\|\nabla_{\mathbf{x}_0} U_k(\mathbf{x}_0^{(j)}, \mathbf{a}^{(j)})\|_2^2$ and $\|\nabla_{\mathbf{a}} U_k(\mathbf{x}_0^{(j)}, \mathbf{a}^{(j)})\|_2^2$.

Hyperparameter reparameterization The posterior energy in Eq. (5) depends explicitly on the variance schedule r_k . We parametrize $r_k = c \sigma_{t_k}$, where σ_{t_k} denotes the diffusion noise level and $c > 0$ is a calibration constant estimated on a healthy validation set to ensure stable reconstruction in the absence of anomalies. As r_k decreases, the relative strength of the likelihood increases compared to the mask prior since it scales as $1/r_k^2$. If the prior weight λ_0 is kept fixed across intermediate steps, this changing scale alters the balance between the likelihood and the logit prior. To preserve a stable trade-off across intermediate steps, we adapt λ_0 . It is chosen such that the gradient magnitudes from the likelihood and mask value are almost equal.

Implementation details We adopt the diffusion parameterization of Karras et al. [10] and train the diffusion model exclusively on healthy data. The conditional estimate $\hat{\mathbf{x}}_0(\mathbf{x}_{t_k}, t_k)$ in Eq. (1) is obtained using the probability flow ordinary differential equation [20] (PF-ODE) with two solver steps per noise level. For all experiments, we use 150 intermediate posterior steps in the DAPS framework and 75 Langevin iterations at each intermediate step. The noise levels $\{t_k\}$ follow

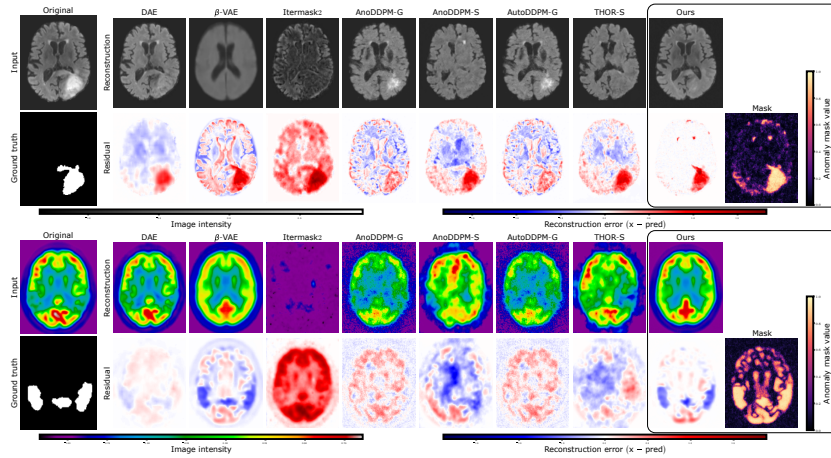


Fig. 2. Qualitative results on BraTS FLAIR (top) and ADNI PET (AD 50%, bottom).

the DAPS parameterization with $\sigma_{\max} = 20$. We set the spatial consistency parameters λ_c to 1. An 80GB NVIDIA H100 GPU was used for training.

4 Experiments & Results

To evaluate the performance of our proposed model, experiments were conducted on 2D FLAIR MRI of glioblastoma patients and on 2D FDG PET with synthetic Alzheimer-like hypometabolism. Our framework was compared with UAD diffusion models [3,4,24] and state of the art methods [9,11,13]. For the proposed approach, we use the inferred latent anomaly mask directly as anomaly maps, while other methods use residuals.

Datasets **ADNI PET** FDG PET images were obtained from ADNI [16] and preprocessed with Clinica [17] (affine registration to MNI, intensity normalization, cropping, resampling to 128^3 , rescaling to $[-1, 1]$). We selected 621 images from 301 cognitively normal subjects. Fifty images were kept for testing. Because no ground truth exists, we use the framework described in [8] to generate synthetic images with Alzheimer-like hypometabolism (50%). The remaining images were split at the subject level into training (545) and validation (25 volumes) sets. **BRATS FLAIR** FLAIR images from BraTS 2021 [15] and ADNI [16] were preprocessed with Clinica [17] in a similar manner as the PET images. Models were trained on healthy ADNI FLAIR slices (234 subjects) and evaluated on BraTS slices containing tumors (1,251 subjects). To analyze sensitivity to anomaly size, slices were grouped into medium and large categories using the 33rd and 66th percentiles of lesion area. One slice per subject was randomly selected per category, yielding 819 large lesions and 1,156 medium lesions.

Evaluation For the PET images, because multiple slices come from the same subject, we account for this hierarchical structure using stratified bootstrapping

Table 1. Performance on BraTS FLAIR. Best possible Dice and AP are reported in % as mean (standard deviation). Arrows indicate favorable direction of each metric.

Category	Method	Large		Medium	
		[Dice] (%) ↑	AP (%) ↑	[Dice] (%) ↑	AP (%) ↑
Reconstruction	DAE [11]	68.6(17.5)	71.9(20.5)	66.9(21.0)	69.0(25.5)
	β -VAE [14]	56.3(16.5)	54.2(22.4)	48.3(19.8)	44.7(25.4)
	IterMask2 [13]	78.7(14.5)	82.3(17.0)	71.0(18.3)	72.8(22.3)
Diffusion	AnoDDPM-G [24]	36.7(9.8)	30.3(11.5)	31.8(13.9)	24.5(15.3)
	AnoDDPM-S [24]	63.1(19.6)	64.9(24.5)	51.5(23.5)	50.1(29.0)
	AutoDDPM-G [3]	44.7(19.4)	39.1(21.5)	45.2(22.7)	40.6(26.6)
	THOR-S [4]	62.6(20.4)	60.4(26.9)	50.3(21.6)	42.5(26.5)
	Ours	72.3(19.2)	73.0(24.3)	63.8(22.6)	62.5(28.3)

Table 2. Performance on ADNI PET. Mean [Dice] and AP are reported in % with 95% bootstrap confidence intervals. Inference speed is reported as slices per second.

Category	Method	Synthetic Alzheimer-like hypometabolism		
		[Dice] (%) ↑	AP (%) ↑	Slices/s ↑
Reconstruction	DAE [11]	47.3[46.0, 48.6]	46.8[44.7, 48.8]	296.7
	β -VAE [9]	61.9[61.2, 62.6]	64.5[63.4, 65.7]	535.4
	IterMask2 [13]	40.1[39.7, 40.6]	17.0[16.7, 17.3]	108.7
Diffusion	AnoDDPM-G [24]	25.2[24.9, 25.4]	14.1[14.0, 14.3]	2.1
	AnoDDPM-S [24]	40.9[39.3, 42.5]	34.7[32.4, 37.3]	2.5
	AutoDDPM-G [3]	9.8[9.3, 10.3]	13.7[13.6, 13.9]	1.2
	THOR-S [4]	43.0[41.6, 44.4]	35.0[33.3, 36.7]	2.1
	Ours	65.7[64.7, 66.7]	67.9[66.2, 69.7]	2.8

at the subject level when computing metrics. We also report the inference time. Detection performance was evaluated using the best possible Dice score ([Dice]) and pixel-wise average precision (AP). Significance was assessed via Wilcoxon signed-rank tests with Bonferroni correction for multiple comparisons ($p < 0.01$).

Baselines We compare our method to diffusion-based UAD approaches, including *AnoDDPM* [24] (G-Gaussian and S-simplex noise variants), *AutoDDPM* (Gaussian noise) [3], and *THOR* (Simplex noise) [4]. For fairness, all diffusion models use the same U-Net backbone. We additionally compare against *IterMask*² [13], a recent state-of-the-art UAD framework that uses synthetic anomalies, and a variational autoencoder (β -VAE) [9,14]. All models are trained on healthy data only. Hyperparameters were tuned on the healthy validation set.

Results Quantitative results on the BraTS dataset are reported in Table 1. Diffusion-based approaches (*AnoDDPM*, *AutoDDPM*, *THOR*) achieve lower AP and [Dice] compared to reconstruction-based methods relying on synthetic anomalies (*IterMask*²) or single-step denoising (*DAE*). Among diffusion models, *Ours*

Table 3. Ablation experiments on BraTS FLAIR (large lesions only). Best possible Dice and AP are reported in % as mean (standard deviation).

Method	[Dice] (%) $\uparrow$	AP (%) $\uparrow$
(A) No latent mask	0.0(0.0)	7.9(2.2)
(B) Fixed likelihood variance	65.8(19.3)	64.6(25.6)
(C) No spatial regularization	46.6(13.9)	38.1(15.4)
Ours	72.3(19.2)	73.0(24.3)

shows the best performance ($p < 0.01$) for both lesion sizes and both evaluation metrics. Compared to all methods, it ranks second on large lesions and third on medium lesions. From Fig. 2, we observe that the *DAE* and *IterMask*² attenuate abnormal hyperintensities toward normal tissue values. In contrast, *Ours* shows restoration of the anomaly, for example reconstructing the ventricles. On ADNI PET (Table 2), our method outperforms all baselines by +3.8pp in [Dice] and +3.4pp in AP ($p < 0.01$). We observe that *IterMask*², which performs strongly on BraTS, fails to generalize to this modality. Finally, inference speed shows that our method does not introduce computational overload compared to standard diffusion models while achieving better detection performance.

Ablation study We evaluate the contribution of each modeling component on the BraTS dataset (Table 3). **(A) No latent mask.** Removing the latent anomaly mask results in no detection; the abnormal image is reconstructed. **(B) Fixed likelihood variance (r_k^2).** Instead of using the variance schedule, the likelihood variance is fixed to 5×10^{-2} as commonly done in classical inverse problems [25]. This modification degrades performance. **(C) No spatial regularization.** Removing spatial consistency in the mask prior (i.e., setting $\lambda_c = 0$) negatively impacts anomaly detection performance. This highlights the importance of spatial regularization to enforce coherent anomaly regions and suppress isolated false positives.

5 Conclusion

In summary, we formulated unsupervised anomaly detection as a Bayesian inverse problem, using a diffusion model as a prior over healthy anatomy. Our method introduces a latent anomaly mask that is inferred jointly with the reconstruction through sampling of a sequence of intermediate posteriors. The mask provides an approximation of the unknown forward model, while a spatial prior ensures coherent and anatomically consistent anomaly regions. Ablation studies confirmed the importance of each proposed component. Experimentally, our model achieves competitive performance on BraTS and state-of-the-art results on ADNI PET compared to multiple baselines.

While the current formulation already demonstrates competitive performance, it also opens several directions for further development. Replacing the Gaussian approximation of the conditional diffusion prior with a more expressive model

could improve the approximation of the intermediate posteriors and enhance anomaly detection. Similarly, refining the forward model formulation may allow the framework to better capture pathological patterns. Since our method relies on joint intermediate posterior inference, alternative sampling or optimization strategies may improve stability and computational efficiency. Extending the framework to full 3D volumes is another important direction. Together, these improvements would extend the framework to other imaging modalities and clinical problems.

Acknowledgments. This work was supported by ANR through the “Investissements d’avenir” program (ANR-10-IAIHU-06, ANR-19-P3IA-0001, PRAIRIE 3IA Institute), the “France 2030” program (ANR-23-IACL-0008, PRAIRIE-PSAI), and the MediTwin project. It also received support from the European Union’s Horizon Europe Framework Programme under grant agreement No. 101136607 (CLARA), and from Next Generation EU through the “France Relance” plan. The ARAMIS Lab is affiliated with DIM C-BRAINS, funded by the Conseil Régional d’Île-de-France. Experiments used GENCI-IDRIS HPC resources under grants 2025-AD011011648R5 and AD011017165. R.D. received Marie Skłodowska-Curie grant No. 101154248 (SafeREG). Data collection and sharing were funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI), supported by NIH grant U19AG024904, with additional public and private-sector support.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Behrendt, F., Bhattacharya, D., Krüger, J., Opfer, R., Schlaefel, A.: Patched Diffusion Models for Unsupervised Anomaly Detection in Brain MRI. In: Medical Imaging with Deep Learning. pp. 1019–1032. PMLR (2024)
2. Beizaee, F., Lodygensky, G., Desrosiers, C., Dolz, J.: Mad-ad: Masked diffusion for unsupervised brain anomaly detection. In: Information Processing in Medical Imaging. pp. 139–153. Springer Nature Switzerland (2026)
3. Bercea, C.I., Neumayr, M., Rueckert, D., Schnabel, J.A.: Mask, Stitch, and Re-Sample: Enhancing Robustness and Generalizability in Anomaly Detection through Automatic Diffusion Models. In: ICML 3rd Workshop on Interpretable Machine Learning in Healthcare (IMLH) (2023)
4. Bercea, C.I., Wiestler, B., Rueckert, D., Schnabel, J.A.: Diffusion Models with Implicit Guidance for Medical Anomaly Detection. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. pp. 211–220. Springer Nature Switzerland, Cham (2024)
5. Cardoso, G., el Idrissi, Y.J., Corff, S.L., Moulines, E.: Monte Carlo guided Denoising Diffusion models for Bayesian linear inverse problems. In: The Twelfth International Conference on Learning Representations (2023)
6. Chung, H., Kim, J., McCann, M.T., Klasky, M.L., Ye, J.C.: Diffusion Posterior Sampling for General Noisy Inverse Problems. In: The Eleventh International Conference on Learning Representations (2022)

7. Dou, Z., Song, Y.: Diffusion posterior sampling for linear inverse problem solving: A filtering perspective. In: International Conference on Learning Representations. vol. 2024, pp. 12069–12095 (2024)
8. Hassanaly, R., Brianceau, C., Solal, M., Colliot, O., Burgos, N.: Evaluation of pseudo-healthy image reconstruction for anomaly detection with deep generative models: Application to brain FDG PET. *Machine Learning for Biomedical Imaging 2*(Special Issue for Generative Models), 611–656 (2024)
9. Hassanaly, R., Solal, M., Colliot, O., Burgos, N., Initiative, f.t.A.D.N.: Benchmarking 3D generative autoencoders for pseudo-healthy reconstruction of brain 18F-fluorodeoxyglucose positron emission tomography. *Journal of Medical Imaging* **12**(5), 054005 (2025)
10. Karras, T., Aittala, M., Aila, T., Laine, S.: Elucidating the Design Space of Diffusion-Based Generative Models. In: *Advances in Neural Information Processing Systems* (2022)
11. Kascenas, A., Pugeault, N., O’Neil, A.Q.: Denoising Autoencoders for Unsupervised Anomaly Detection in Brain MRI. In: *Proceedings of The 5th International Conference on Medical Imaging with Deep Learning*. pp. 653–664. PMLR (2022)
12. Kavar, B., Elad, M., Ermon, S., Song, J.: Denoising diffusion restoration models. In: *Advances in Neural Information Processing Systems* (2022)
13. Liang, Z., Guo, X., Noble, J.A., Kamnitsas, K.: Itermask2: Iterative Unsupervised Anomaly Segmentation via Spatial and Frequency Masking for Brain Lesions in MRI. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*. pp. 339–348. Springer Nature Switzerland, Cham (2024)
14. Loizillon, S., Jacob, Y., Aurélien, M., Dormont, D., Colliot, O., Burgos, N., Group, A.S.: Detecting Brain Anomalies in Clinical Routine with the β -VAE: Feasibility Study on Age-Related White Matter Hyperintensities. In: *Proceedings of The 7nd International Conference on Medical Imaging with Deep Learning*. pp. 903–917. PMLR (2024)
15. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M.A., Arbel, T., Avants, B.B., Ayache, N., Buendia, P., Collins, D.L., Cordier, N., Corso, J.J., Criminisi, A., Das, T., Delingette, H., Demiralp, Ç., Durst, C.R., Dojat, M., Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Guo, X., Hamamci, A., Iftekharuddin, K.M., Jena, R., John, N.M., Konukoglu, E., Lashkari, D., Mariz, J.A., Meier, R., Pereira, S., Precup, D., Price, S.J., Raviv, T.R., Reza, S.M.S., Ryan, M., Sarikaya, D., Schwartz, L., Shin, H.C., Shotton, J., Silva, C.A., Sousa, N., Subbanna, N.K., Szekely, G., Taylor, T.J., Thomas, O.M., Tustison, N.J., Unal, G., Vasseur, F., Wintermark, M., Ye, D.H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., Van Leemput, K.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging* **34**(10), 1993–2024 (2015)
16. Petersen, R.C., Aisen, P.S., Beckett, L.A., Donohue, M.C., Gamst, A.C., Harvey, D.J., Jack, C.R., Jagust, W.J., Shaw, L.M., Toga, A.W., Trojanowski, J.Q., Weiner, M.W.: Alzheimer’s Disease Neuroimaging Initiative (ADNI): Clinical characterization. *Neurology* **74**(3), 201–209 (2010)
17. Routier, A., Burgos, N., Díaz, M., Bacci, M., Bottani, S., El-Rifai, O., Fontanella, S., Gori, P., Guillon, J., Guyot, A., Hassanaly, R., Jacquemont, T., Lu, P., Marcoux, A., Moreau, T., Samper-González, J., Teichmann, M., Thibeau-Sutre, E., Vaillant, G., Wen, J., Wild, A., Habert, M.O., Durrleman, S., Colliot, O.: Clinica: An Open-

- Source Software Platform for Reproducible Clinical Neuroscience Studies. *Frontiers in Neuroinformatics* **15**, 39 (2021)
18. Roy, H., Dorent, R., Burgos, N.: Unsupervised anomaly detection using Bayesian flow networks: application to brain FDG pet in the context of Alzheimer’s disease. In: *Lecture Notes in Computer Science : Deep Generative Models* (2025)
 19. Song, J., Vahdat, A., Mardani, M., Kautz, J.: Pseudoinverse-Guided Diffusion Models for Inverse Problems. In: *International Conference on Learning Representations* (2022)
 20. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B.: Score-Based Generative Modeling through Stochastic Differential Equations. In: *International Conference on Learning Representations* (2020)
 21. Welling, M., Teh, Y.W.: Bayesian learning via stochastic gradient langevin dynamics. In: *Proceedings of the 28th International Conference on International Conference on Machine Learning* (2011)
 22. Wolleb, J., Bieder, F., Friedrich, P., Zhang, P., Durrer, A., Cattin, P.C.: Binary Noise for Binary Tasks: Masked Bernoulli Diffusion for Unsupervised Anomaly Detection. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*. pp. 135–145. Springer Nature Switzerland, Cham (2024)
 23. Wolleb, J., Bieder, F., Sandkühler, R., Cattin, P.C.: Diffusion Models for Medical Anomaly Detection. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. pp. 35–45. Springer Nature Switzerland, Cham (2022)
 24. Wyatt, J., Leach, A., Schmon, S.M., Willcocks, C.G.: AnoDDPM: Anomaly Detection With Denoising Diffusion Probabilistic Models Using Simplex Noise. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022)
 25. Zhang, B., Chu, W., Berner, J., Meng, C., Anandkumar, A., Song, Y.: Improving Diffusion Inverse Problem Solving with Decoupled Noise Annealing. In: *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 20895–20905 (2025)