

REFLECT: Rectified Flows for Efficient Brain Anomaly Correction Transport.

Farzad Beizaee^{1,2}, Sina Hajimiri¹, Ismail Ben Ayed¹,
Gregory Lodygensky², Christian Desrosiers¹, and Jose Dolz¹

¹ ÉTS Montreal, Canada

² CHU Sainte-Justine, University of Montreal, Canada
`farzad.beizaee.1@ens.etsmtl.ca`

Abstract. Unsupervised anomaly detection (UAD) in brain imaging is crucial for identifying pathologies without the need for labeled data. However, accurately localizing anomalies remains challenging due to the intricate structure of brain anatomy and the scarcity of abnormal examples. In this work, we introduce REFLECT, a novel framework that leverages rectified flows to establish a direct, linear trajectory for correcting abnormal MR images toward a normal distribution. By learning a straight, one-step correction transport map, our method efficiently corrects brain anomalies and can precisely localize anomalies by detecting discrepancies between anomalous input and corrected counterpart. In contrast to the diffusion-based UAD models, which require iterative stochastic sampling, rectified flows provide a direct transport map, enabling single-step inference. Extensive experiments on popular UAD brain segmentation benchmarks demonstrate that REFLECT significantly outperforms state-of-the-art unsupervised anomaly detection methods. The code is available at <https://github.com/farzad-bz/REFLECT>.

Keywords: Rectified flows · Unsupervised anomaly detection · Brain MRI

1 Introduction

Brain anomaly detection using medical images is a critical task in neuroimaging, with significant implications for early diagnosis and treatment planning. Brain abnormalities such as tumors, lesions, or traumatic injuries often appear as structural deviations from normal anatomy. While supervised methods for localizing anomalies are effective in many settings, they rely on large annotated datasets, which are costly and often scarce, particularly for rare anomalies. This has led to growing interest in unsupervised methods, where models learn the normal data distribution and subsequently localize deviations as potential abnormalities.

In recent years, generative models have been widely investigated for reconstruction-based unsupervised anomaly detection (UAD). Initial attempts leverage Auto-Encoders (AEs) [2,3] and their variants, including Variational Auto-Encoders (VAEs) [28,36]. Generative Adversarial Networks (GANs) [13],

including AnoGAN [27] and f-AnoGAN [26], have also emerged as promising alternatives to AE-based methods. However, these approaches tend to overfit the normal training data or yield blurry reconstructions. Normalizing flows (NFs) [31,24] are another family of generative models that transform a simple base distribution into a complex target distribution through a series of invertible transformations, thereby allowing exact likelihood estimation. This makes them particularly appealing for anomaly localization via out-of-distribution detection [14,17,10,35]. However, NFs require complex architectures, are computationally expensive, and often involve iterative steps.

Progress in generative modeling has led to the rise of diffusion models [15], a powerful class of probabilistic models that generate high-quality data by gradually transforming noise into structured outputs through a learned iterative process. Fueled by their impressive generative performance, diffusion models have been increasingly adopted in medical imaging tasks, including unsupervised anomaly detection [4,5,6,8,19,22,23,34]. AnoDDPM [34] resorts to a partial diffusion strategy, where it adds noise to the image up to a certain timestep and then recovers it via reverse diffusion, whereas pDDPM [4] applies diffusion patch-wise to better capture local context. THOR [8] refines diffusion models by using implicit temporal guidance via anomaly maps during the reconstruction process. Inspired by diffusion models, Itermask [19] proposes iterative mask refinement using reconstruction errors to better localize brain lesions in MRI. And very recently, MAD-AD [6] treats abnormalities as noise in the latent space and uses a masked diffusion process to selectively correct abnormalities. While diffusion models offer a better overall performance, they tend to “memorize” patterns from training data [29], which reduces their generalizability. Also, they need many iterative steps to reconstruct normal images, even with DDIM sampling [30]. Last but not least, all these generative methods are primarily designed to generate new samples (often from pure noise), and not to modify existing images. This limits their applicability for the selective correction of images.

An alternative unsupervised approach uses self-supervised learning to restore normal images corrupted with synthetic anomalies. Methods like Foreign Patch Interpolation (FPI)[32] and Poisson Image Interpolation (PII)[33] introduce such defects by blending patches from different normal images, enabling pixel-level anomaly localization. Most recently, DISYRE v2 [23] introduced a cold-diffusion pipeline that restores synthetically corrupted images with controlled anomaly severity through iterative refinement. While these methods are effective, their reliance on synthetic anomalies limits generalization, as these may not fully capture the variability of real-world abnormalities.

To address the aforementioned limitations, we propose REFLECT, an unsupervised brain anomaly detection framework built on the recently introduced rectified flows [21]. Rectified flows, which learn a transport map between two different distributions through rectified trajectories, exhibit several advantages over the previously mentioned generative models, such as improved stability, high-fidelity reconstructions, and direct and efficient mapping, which enables the mapping of a sample to a target distribution with a single step.

Our work makes the following key contributions: *i)* We propose leveraging rectified flows in latent space to enable optimal and straight transport of abnormal brain samples toward their normal counterparts, thereby facilitating unsupervised brain anomaly detection with enhanced accuracy and reliability. To the best of our knowledge, rectified flow has not been explored before for unsupervised anomaly detection. By learning straight flow trajectories, our approach requires fewer time steps, which enables high-quality correction of abnormal regions in a single step while preserving normal regions. Unlike the often complex formulations used in diffusion models or GANs, rectified flows’ objective directly controls the geometry of the flow and provides a natural mechanism to correct anomalous samples, making the method both theoretically elegant and practically efficient. *ii)* Moreover, we introduce an effective technique for generating diverse and realistic anomalous brain MRIs using a random walk-based masking strategy operating in the latent space. Masked regions are replaced with textured image segments or random noise, enhancing the variability and realism of synthesized anomalies. *iii)* Extensive experiments on brain anomaly detection benchmarks further demonstrate the superiority of our approach over recent state-of-the-art unsupervised anomaly detection methods.

2 Preliminaries: Rectified Flows

Rectified flows [21], building upon the principles of continuous normalizing flows [9], learn an ordinary differential equation (ODE) that transports samples from an initial distribution π_0 to a target distribution π_1 along trajectories that are as straight as possible in a continuous-time framework. Concretely, consider the linear interpolation between random variables $X_0 \sim \pi_0$ and $X_1 \sim \pi_1$:

$$X_t = (1 - t) X_0 + t X_1, \quad t \in [0, 1]. \quad (1)$$

Although X_t provides a continuous trajectory from X_0 to X_1 , these straight-line paths are non-causal and may intersect when considering different sample pairs. Such intersections are undesirable for generative modeling, as they lead to non-causal and non-deterministic dynamics. This interpolation is “rectified” into a causal ODE flow $\{Z_t : t \in [0, 1]\}$ via

$$\frac{dZ_t}{dt} = v(Z_t, t), \quad Z_0 \sim \pi_0, \quad (2)$$

where Z_t represents the state of the rectified flow at time t , and $v(\cdot, t)$ is a trainable velocity field. We seek v that best aligns with the direction of linear interpolation paths. Therefore, the training objective minimizes the discrepancy between the true instantaneous velocity of the interpolation and the learned velocity:

$$\min_v \int_0^1 \mathbb{E} \left[\|(X_1 - X_0) - v_\theta(X_t, t)\|^2 \right] dt. \quad (3)$$

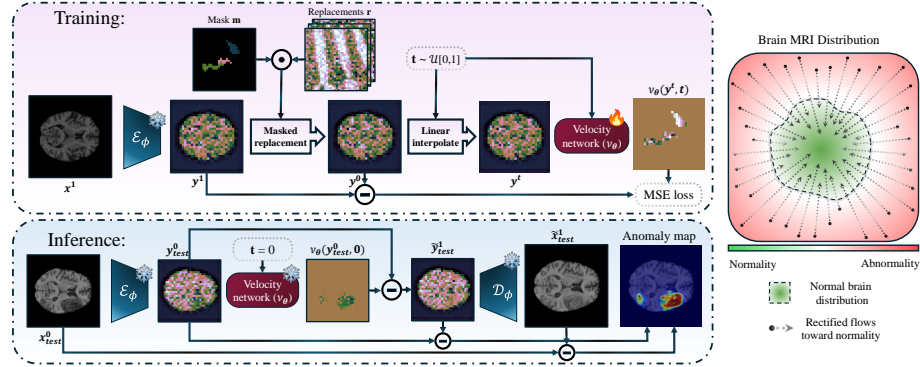


Fig. 1. Overview of REFLECT. *Training:* a velocity network is trained to predict the displacement between the latent representations of normal images and their corrupted versions, using their linear interpolation and time t . *Inference:* Given an anomalous test image, the velocity network computes the flow, allowing the image to be corrected in a single step. *Right:* Rectified flows toward normality in the brain’s latent space.

In practice, we approximate the integral via Monte Carlo sampling:

$$\min_{\theta} \mathbb{E}_{t \sim U[0,1]} \left[\left\| (X_1 - X_0) - v_{\theta}(X_t, t) \right\|^2 \right]. \quad (4)$$

Moreover, the *reflow* procedure, an iterative application of rectified flow training to the outputs of a previous flow, can further straighten trajectories and minimize discretization errors. This process helps prevent trajectory crossings induced by paired input samples and further refines and rectifies the transport paths toward the target distribution, thereby enhancing robustness and enabling high-quality single-step sampling while preserving distribution fidelity. This is achieved through the following objective:

$$\min_{\theta} \mathbb{E}_{Z_0 \sim \pi_0, t \sim U(0,1)} \left[\left\| (Z_1 - Z_0) - v'_{\theta}(Z_t, t) \right\|^2 \right], \quad Z_0 = X_0, \quad (5)$$

where Z_1 represents the transported sample (Z_0) with the initial rectified flow.

3 Method

Interpreting π_0 as the entire brain distribution encompassing both normal and abnormal (or injured) brains, and π_1 as the distribution of normal brains (with $\pi_1 \subset \pi_0$), we train a rectified flow to learn a velocity field $v_{\theta}(\cdot, t)$ that transports abnormal samples from π_0 toward π_1 (illustrated in Fig. 1), while leaving normal samples unchanged since they already belong to π_1 . Our framework for unsupervised brain anomaly detection using rectified flows (see Figure 1) consists of three main stages: (1) generating paired abnormal and normal images in the latent space, (2) rectified flow training to transport abnormal brains toward normality, and (3) localizing anomalies. Below, we describe each stage in detail.

3.1 Generating Paired Samples

Following [12,6], we transform the data into a latent space to stabilize training, improve output quality, and enable the interpretation of anomalies as noise or out-of-distribution features. To achieve this, we adapt and fine-tune a pre-trained VAE [25] for medical images, which compresses high-dimensional data into a compact latent representation while retaining structures and semantic information. Let $\mathcal{X} = \{\mathbf{x}^{(i)}\}_{i=1}^N$ denote the set of normal brain samples; the encoder \mathcal{E}_ϕ maps each image to a latent code via $\mathbf{y}^{(i)} = \mathcal{E}_\phi(\mathbf{x}^{(i)})$.

To train the rectified flow for transforming anomalous latent representations to normal ones, we construct paired latent vectors $(\mathbf{y}_i^0, \mathbf{y}_i^1)$, where \mathbf{y}_i^1 is a normal sample and \mathbf{y}_i^0 is an artificially corrupted version. The corruption is applied using a binary mask \mathbf{m} that specifies the regions to be altered, a random replacement vector \mathbf{r} indicating the direction of corruption, and parameter $\alpha \in [0, 1]$ which controls the severity of corruption. The corrupted latent vector is given by:

$$\mathbf{y}_i^0 = (\sqrt{1-\alpha} \cdot \mathbf{y}_i^1 + \sqrt{\alpha} \cdot \mathbf{r}) \odot \mathbf{m} + \mathbf{y}_i^1 \odot (1 - \mathbf{m}), \quad (6)$$

which guarantees that if both \mathbf{y}_i^1 and \mathbf{r} follow a standard normal distribution, the resulting corrupted image will also follow a standard normal distribution.

The introduced corruptions must closely mimic genuine injuries and span the entire range of possible anomalies. We assume that the brain may exhibit up to N distinct anomalous regions delineated by non-overlapping masks. The mask of each region is generated by a random walk starting from a random point within the brain and taking a random number of uniformly probable steps to neighboring points. Masked positions then correspond to the points visited during the walk. This strategy results in more realistic anomaly shapes than rectangular patch-based masking [6].

Afterward, a replacement vector is assigned to each masked region, and the final corrupted image is generated according to Eq. 6. For each region, the replacement vector is generated by randomly choosing between two distinct strategies. The first strategy is to use random noise, motivated by the interpretation of anomalies as noise in the latent space. Alternatively, we can use a cropped segment from the latent representation of a textured image (not necessarily medical images) to impose more realistic and structured anomalies. For the first strategy, in our implementation, we propose sampling random noise independently for each channel according to:

$$\mathbf{r} = \sqrt{\beta} \cdot \mathbf{q} + \sqrt{1-\beta} \cdot \mathbf{p}, \quad (7)$$

where $\mathbf{q} \in \mathbb{R} \sim \mathcal{N}(0, 1)$, $\mathbf{p} \in \mathbb{R}^{H \times W} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and β is uniformly sampled from $[0, 1]$ (fixed for all channels). This formulation can be interpreted as a weighted combination of a global (image-level) random variable and a spatially varying (pixel-wise) random vector, constructed such that the resulting \mathbf{r} follows a standard normal distribution. This strategy imposes spatial dependencies and yields more realistic and structured corruptions in the latent space.

3.2 Training Rectified Flows for Anomaly Correction

Our goal is to learn a mapping that transports abnormal latent samples $\mathbf{y}^0 \sim \pi_0$ back to the normal distribution π_1 . For this purpose, we train the velocity field $v_\theta(\cdot, t)$ using the paired latent data $(\mathbf{y}^0, \mathbf{y}^1)$. The training objective becomes:

$$\min_{\theta} \mathbb{E}_{t \sim \mathcal{U}[0,1]} \left[\|(Y_1 - Y_0) - v_\theta(Y_t, t)\|^2 \right], \quad (8)$$

where $Y_t = (1-t)Y_0 + tY_1$. Once the velocity model is trained, the learned ODE (Eq. 2) defines a flow that can be numerically simulated using a standard ODE solver (e.g., Euler method). Note that in many cases, including ours, a coarse discretization (even a single Euler step) is sufficient due to the straightening effect of the learned flow. Furthermore, after obtaining the first velocity model, we can refine the flow by training a second rectified flow model $v'(\cdot, t)$ using the reflow process. This additional stage could further straighten the flows and enhance the correction trajectories toward the target normal distribution (i.e., anomaly-free brains). We refer to this model as 2-REFLECT, while the first model is denoted as 1-REFLECT.

3.3 Inference and Anomaly Localization

At test time, given an image $\mathbf{x}_{\text{test}}^0$ (potentially containing anomalies), we first encode it as $\mathbf{y}_{\text{test}}^0 = \mathcal{E}_\phi(\mathbf{x}_{\text{test}}^0)$. Then, we solve the reverse-time ODE with the trained velocity model using a single Euler step, which results in:

$$\tilde{\mathbf{y}}_{\text{test}}^1 = \mathbf{y}_{\text{test}}^0 - v_\theta(\mathbf{y}_{\text{test}}^0, 0), \quad (9)$$

where $\tilde{\mathbf{y}}_{\text{test}}^1$ should ideally lie within the π_1 distribution. The reconstructed normal image is decoded from the corrected latent space using the VAE decoder: $\tilde{\mathbf{x}}_{\text{test}}^1 = \mathcal{D}_\phi(\tilde{\mathbf{y}}_{\text{test}}^1)$. Finally, anomalies are localized by comparing the original image to its reconstruction in both latent space and image space:

$$\mathbf{A}(\mathbf{x}_{\text{test}}) = \frac{1}{2} |\tilde{\mathbf{x}}_{\text{test}}^1 - \mathbf{x}_{\text{test}}^0| + \frac{1}{2} |\tilde{\mathbf{y}}_{\text{test}}^1 - \mathbf{y}_{\text{test}}^0|. \quad (10)$$

Higher differences indicate regions where the flow had to make larger corrections, thereby signaling anomalies.

4 Experiments

4.1 Experimental Setting

Datasets. For our experiments, we employed BraTS'21 [1] that comprises 1251 brain scans across four modalities (T1, Contrast-Enhanced T1 (T1CE), T2, and FLAIR), and ATLAS 2.0 [20], which contains 655 T1-weighted MRI scans. Both datasets come with expert-annotated lesion masks. We extracted 20 central axial slices from skull-stripped brains and padded them to a resolution of 256×256

Table 1. Quantitative results obtained by different approaches. The best method per modality and/or dataset is highlighted in **bold**, whereas the second one is underlined. The performance gains over the best baseline are shown in **green**.

Method		ATLAS 2.0	BraTS'21			
		T1-w	FLAIR	T1CE	T2-w	T1-w
AE [2]	<i>MedIA'21</i>	11.9	33.4	32.3	30.2	28.5
DDPM [15]	<i>Neurips'20</i>	20.2	60.7	37.9	36.4	29.4
AutoDDPM [7]	<i>ICML Workshop'23</i>	12.7	55.5	36.9	29.7	33.5
DAE [16]	<i>MedIA'22</i>	11.1	79.7	36.7	69.6	29.5
Cycl.UNet [18]	<i>MICCAI'23</i>	N/A	65.0	42.6	49.5	37.0
IterMask ² [19]	<i>MICCAI'24</i>	35.3	<u>80.2</u>	61.7	71.2	58.5
MAD-AD [6]	<i>IPMI'25</i>	<u>36.1</u>	<u>76.2</u>	<u>68.5</u>	<u>73.2</u>	<u>63.4</u>
1-REFLECT	<i>Ours</i>	41.6 ± 5.5	85.1 ± 4.9	73.0 ± 4.5	79.6 ± 6.4	69.7 ± 6.3
2-REFLECT	<i>Ours</i>	40.8 ± 4.7	83.2 ± 3.0	72.0 ± 3.5	80.3 ± 7.1	69.8 ± 6.4

pixels. Both datasets are divided into training (80%), validation (10%), and testing (10%) subsets, and only normal training slices are used for training, while the single slice of the test subjects displaying the most prominent pathology is reserved for inference. Moreover, the Describable Textures Dataset (DTD) [11] is used and converted to gray-scale to serve as textured replacement images.

Evaluation metrics. Following [6], we assess models' performance using the Maximum Dice score, which reflects the highest Dice coefficient achieved as the threshold varies from 0 to 1.

Implementation details. In each iteration, the number of masked regions is randomly chosen between 1 and 4, and the number of random-walk steps is randomly sampled between 0 and 200. 1-REFLECT model underwent training for 200 epochs for BraTS'21 and for 400 epochs for the ATLAS dataset with a batch size of 96, using the AdamW optimizer and a learning rate of 5×10^{-4} . Afterward, 2-REFLECT was trained on top of 1-REFLECT for another 50 epochs with a learning rate of 1×10^{-5} . Also, we have used 5 reverse ODE correction steps.

4.2 Results

Main quantitative results. We empirically evaluate the performance of the proposed approach compared to relevant UAD methods proposed for brain MRI, whose results are reported in Table 1. These values demonstrate that REFLECT substantially outperforms recent state-of-the-art brain UAD methods. Compared to the second-best approach, i.e., MAD-AD [6], our method brings improvement gains ranging from 4.5% to 6.4% in both datasets. In particular, in T1-w, which seems to be the most difficult modality based on baselines' results, REFLECT yields the highest difference gaps (similar to gains in T2-w). These results demonstrate the superiority of our method across modalities and datasets, compared to very relevant baselines. Regarding **reflow's** effect (i.e., 2-REFLECT) on the first trained model 1-REFLECT, results (Table 1) suggest that the reflow process

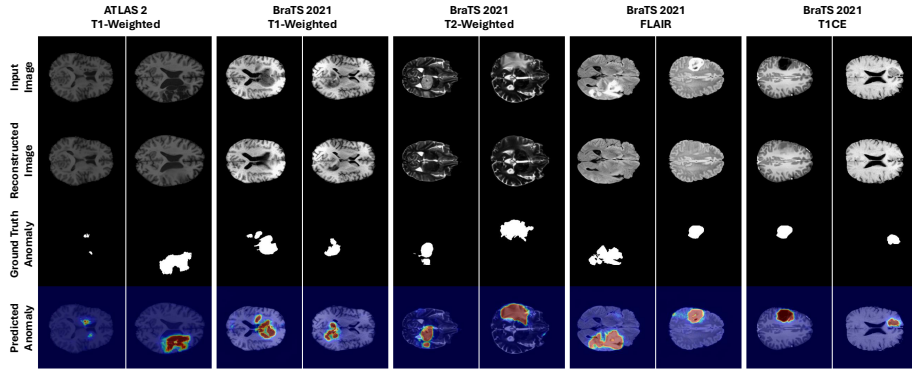


Fig. 2. Qualitative results. *Second row:* Reconstructed images of their abnormal input counterparts. *Last row:* Anomaly segmentation maps obtained by our approach.

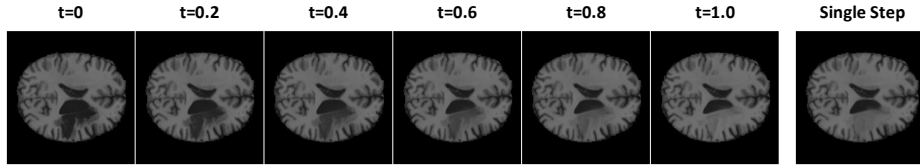


Fig. 3. Transition between an anomalous to healthy brain using 10 reverse ODE steps *vs.* a single step, showing that flows are well-rectified and effective in a single step.

did not improve the overall performance of the anomaly detection, likely because the initial flows are already well-rectified, and the reflow process is attached to the anomaly localization performance of the first method.

Qualitative results. To visually assess the effectiveness of REFLECT, we depict in Fig. 2 several visual examples of both the reconstructed images (*second row*) and the predicted anomaly maps (*last row*) across all modalities of the BraTS dataset. As these images highlight, REFLECT successfully reconstructs realistic “healthy” images, leading to accurately predicted anomaly maps. Fig. 3 depicts a visual example of the transitions performed by the rectified flow.

Ablations. *Model size:* We evaluated five different model sizes to analyze the trade-off between model efficiency and anomaly detection performance. While larger models demonstrated slightly better performance, smaller variants still achieved highly competitive results, significantly outperforming previous methods, making REFLECT a strong choice for real-time or resource-constrained applications. *Effect of the VAE model:* We also investigated the effect of the VAE model by comparing two variants with scale factors of 4 and 8. As illustrated in Table 2, both VAEs perform well, with the VAE using a scale factor of 4 exhibiting slightly better performance, likely due to its higher embedding dimensionality. However, this comes at the cost of increased computational requirements.

Table 2. Ablation studies (BRATS’21) on model size and VAE model employed.

	Model	#Params	Modality				
			FLAIR	T1CE	T2-w	T1-w	Avg
Model size	UNet XS	~16 M	81.7	70.6	80.9	69.7	75.7
	UNet S	~64 M	83.9	72.1	81.0	69.6	76.7
	UNet M*	~145 M	85.1	73.0	79.6	69.7	76.9
	UNet L	~257 M	83.2	72.1	80.8	71.2	76.8
	UNet XL	~580 M	83.6	73.3	79.9	70.7	76.9
VAE	KL-f4	~55 M	85.4	72.1	80.8	71.8	77.5
	KL-f8*	~84 M	85.1	73.0	79.6	69.7	76.9

5 Conclusion

We introduced REFLECT, an unsupervised framework that leverages rectified flows in latent space, which provides a direct one-step correction transport that significantly improves reconstruction fidelity and anomaly localization. Experimental results on established unsupervised anomaly detection benchmarks confirm that REFLECT outperforms current state-of-the-art approaches, paving the way for more robust and efficient diagnostic tools in neuroimaging. Future directions include extending the framework to 3D data and testing cross-dataset generalization to boost clinical relevance.

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

References

1. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J.S., Freymann, J.B., Farahani, K., Davatzikos, C.: Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data* **4**(1), 1–13 (2017)
2. Baur, C., Denner, S., Wiestler, B., Navab, N., Albarqouni, S.: Autoencoders for unsupervised anomaly segmentation in brain MR images: a comparative study. *Medical Image Analysis* **69**, 101952 (2021)
3. Baur, C., Wiestler, B., Albarqouni, S., Navab, N.: Deep autoencoding models for unsupervised anomaly segmentation in brain MR images. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018*. pp. 161–169 (2019)
4. Behrendt, F., Bhattacharya, D., Krüger, J., Opfer, R., Schlaefer, A.: Patched diffusion models for unsupervised anomaly detection in brain MRI. In: *Medical Imaging with Deep Learning*. pp. 1019–1032. PMLR (2024)
5. Behrendt, F., Bhattacharya, D., Mieling, R., Maack, L., Krüger, J., Opfer, R., Schlaefer, A.: Leveraging the Mahalanobis distance to enhance unsupervised brain MRI anomaly detection. In: *MICCAI*. pp. 394–404 (2024)
6. Beizae, F., Lodygensky, F., Desrosiers, C., Dolz, J.: MAD-AD: Masked diffusion for unsupervised brain anomaly detection. In: *IPMI* (2025)

7. 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
8. Bercea, C.I., Wiestler, B., Rueckert, D., Schnabel, J.A.: Diffusion models with implicit guidance for medical anomaly detection. In: MICCAI. pp. 211–220 (2024)
9. Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations. In: NeurIPS. pp. 6571–6583 (2018)
10. Chiu, L.L., Lai, S.H.: Self-supervised normalizing flows for image anomaly detection and localization. In: CVPR. pp. 2927–2936 (2023)
11. Cimpoi, M., Maji, S., Kokkinos, I., Mohamed, S., Vedaldi, A.: Describing textures in the wild. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3606–3613 (2014)
12. Esser, P., Kulal, S., Blattmann, A., Entezari, R., Müller, J., Saini, H., Levi, Y., Lorenz, D., Sauer, A., Boesel, F., et al.: Scaling rectified flow transformers for high-resolution image synthesis. In: ICML (2024)
13. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. In: NeurIPS. pp. 2672–2680 (2014)
14. Gudovskiy, D., Ishizaka, S., Kozuka, K.: Cflow-ad: Real-time unsupervised anomaly detection with localization via conditional normalizing flows. In: WACV. pp. 98–107 (2022)
15. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. *NeurIPS* **33**, 6840–6851 (2020)
16. Kascenas, A., Pugeault, N., O’Neil, A.Q.: Denoising autoencoders for unsupervised anomaly detection in brain mri. In: International Conference on Medical Imaging with Deep Learning. pp. 653–664. PMLR (2022)
17. Kim, D., Baik, S., Kim, T.H.: Sanflow: Semantic-aware normalizing flow for anomaly detection. *NeurIPS* **36**, 75434–75454 (2023)
18. Liang, Z., Anthony, H., Wagner, F., Kamnitsas, K.: Modality cycles with masked conditional diffusion for unsupervised anomaly segmentation in MRI. In: MICCAI. pp. 168–181 (2023)
19. Liang, Z., Guo, X., Noble, J.A., Kamnitsas, K.: Itermask 2: Iterative unsupervised anomaly segmentation via spatial and frequency masking for brain lesions in MRI. In: MICCAI. pp. 339–348 (2024)
20. Liew, S.L., Lo, B.P., Donnelly, M.R., Zavaliangos-Petropulu, A., Jeong, J.N., Barisano, G., Hutton, A., Simon, J.P., Juliano, J.M., Suri, A., et al.: A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms. *Scientific data* **9**(1), 320 (2022)
21. Liu, X., Gong, C., Liu, Q.: Flow straight and fast: Learning to generate and transfer data with rectified flow. In: The Eleventh International Conference on Learning Representations (ICLR) (2023)
22. Marimont, S.N., Baugh, M., Siomos, V., Tzelepis, C., Kainz, B., Tarroni, G.: Disyre: Diffusion-inspired synthetic restoration for unsupervised anomaly detection. In: 2024 IEEE International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2024)
23. Marimont, S.N., Siomos, V., Baugh, M., Tzelepis, C., Kainz, B., Tarroni, G.: Ensembled cold-diffusion restorations for unsupervised anomaly detection. In: MICCAI. pp. 243–253 (2024)
24. Rezende, D.J., Mohamed, S.: Variational inference with normalizing flows. In: ICML. pp. 1530–1538 (2015)

25. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)
26. Schlegl, T., Seeböck, P., Waldstein, S.M., Langs, G., Schmidt-Erfurth, U.: f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Medical image analysis* **54**, 30–44 (2019)
27. Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G.: Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: IPMI. pp. 146–157 (2017)
28. Silva-Rodríguez, J., Naranjo, V., Dolz, J.: Constrained unsupervised anomaly segmentation. *Medical Image Analysis* **80**, 102526 (2022)
29. Somepalli, G., Singla, V., Goldblum, M., Geiping, J., Goldstein, T.: Understanding and mitigating copying in diffusion models. *Advances in Neural Information Processing Systems* **36**, 47783–47803 (2023)
30. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: International Conference on Learning Representations (2021)
31. Tabak, E.G., Turner, C.V.: A family of nonparametric density estimation algorithms. *Communications on Pure and Applied Mathematics* **66**(2), 145–164 (2013)
32. Tan, J., Hou, B., Batten, J., Qiu, H., Kainz, B., et al.: Detecting outliers with foreign patch interpolation. *Machine Learning for Biomedical Imaging* **1**(April 2022 issue), 1–27 (2022)
33. Tan, J., Hou, B., Day, T., Simpson, J., Rueckert, D., Kainz, B.: Detecting outliers with poisson image interpolation. In: MICCAI. pp. 581–591. Springer (2021)
34. Wyatt, J., Leach, A., Schmon, S.M., Willcocks, C.G.: ANODDPM: Anomaly detection with denoising diffusion probabilistic models using simplex noise. In: CVPRw. pp. 650–656 (2022)
35. Zhao, Y., Ding, Q., Zhang, X.: AE-FLOW: Autoencoders with normalizing flows for medical images anomaly detection. In: ICLR (2023)
36. Zimmerer, D., Isensee, F., Petersen, J., Kohl, S., Maier-Hein, K.: Unsupervised anomaly localization using variational auto-encoders. In: MICCAI. pp. 289–297 (2019)