



What Do AEs Learn? Challenging Common Assumptions in Unsupervised Anomaly Detection

Cosmin I. Bercea^{1,2(✉)}, Daniel Rueckert^{1,3}, and Julia A. Schnabel^{1,2,4}

¹ Technical University of Munich, Munich, Germany
cosmin.bercea@tum.de

² Helmholtz AI and Helmholtz Center Munich, Munich, Germany

³ Imperial College London, London, UK

⁴ King's College London, London, UK

Abstract. Detecting abnormal findings in medical images is a critical task that enables timely diagnoses, effective screening, and urgent case prioritization. Autoencoders (AEs) have emerged as a popular choice for anomaly detection and have achieved state-of-the-art (SOTA) performance in detecting pathology. However, their effectiveness is often hindered by the assumption that the learned manifold only contains information that is important for describing samples within the training distribution. In this work, we challenge this assumption and investigate what AEs actually learn when they are posed to solve anomaly detection tasks. We have found that standard, variational, and recent adversarial AEs are generally not well-suited for pathology detection tasks where the distributions of normal and abnormal strongly overlap. In this work, we propose *MorphAEus*, novel deformable AEs to produce pseudo-healthy reconstructions refined by estimated dense deformation fields. Our approach improves the learned representations, leading to more accurate reconstructions, reduced false positives and precise localization of pathology. We extensively validate our method on two public datasets and demonstrate SOTA performance in detecting pneumonia and COVID-19. Code: <https://github.com/ci-ber/MorphAEus>.

Keywords: Representation Learning · Anomaly Detection

1 Introduction

Identifying unusual patterns in data is of great interest in many applications such as medical diagnosis, industrial defect inspection, or financial fraud detection. Finding anomalies in medical images is especially hard due to large inter-patient variance of normality, the irregular appearance-, and often rare occurrence of diseases. Therefore, it is difficult and expensive to collect large amounts of annotated samples that cover the full abnormality spectrum, with supervised [7, 15]

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43904-9_30.

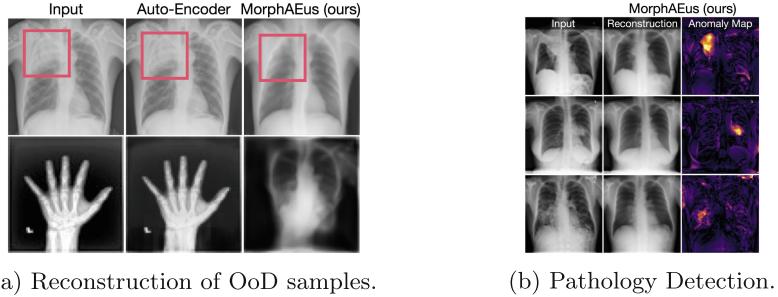


Fig. 1. *MorphAEus* outperforms AEs by generating ID reconstructions even for far OoD cases (Fig. 1a), enabling accurate pathology localization (Figure 1b).

and self-supervised [10, 16] methods only capturing limited facets of the abnormal distribution [31]. However, since it is more feasible to obtain large data sets with normal samples, it is common to detect outliers by detecting patterns that deviate from the expected normative distribution.

Reconstruction-based AEs have emerged as a very popular framework for unsupervised anomaly detection and are widely adopted in medical imaging [2]. They provide straight-forward residual error maps, which are essential for safety-critical domains such as medical image analysis. However, recent work suggests that AEs might reconstruct out-of-distribution (OoD) samples even better than in-distribution (ID) samples [28], with the learned likelihood being dominated by common low-level features [32]. While this can be useful for some tasks such as reconstruction [34], or restoration [21], it often fails for pathology detection as anomalies can be missed due to small residual errors. In Fig. 1 we show that AEs that have only been trained on healthy chest X-rays are also able to reconstruct OoD samples like pathologies or hands. Similarly, Perera et al. [24] showed that AEs trained on the digit 8 can also reconstruct digits 1,5,6 and 9.

Much effort has been made in the medical imaging community to improve the limitations of traditional anomaly detection methods, particularly in the context of brain MRI. Apart from the reduced dimensionality of the bottleneck, several other techniques have been introduced to regularize AEs [11, 20, 29, 38]. Recently, self-supervised denoising AEs [16] achieved SOTA results on brain pathology segmentation. They explicitly feed noise-corrupted inputs $\tilde{x} = x + \epsilon$ to the network with the aim at reconstructing the original input x . However, this is especially beneficial when the anomaly distribution is known *a priori*. Variational AEs (VAEs) [5, 12, 17, 40] estimate the distribution over the latent space that is regularized to be similar to a prior distribution, usually a standard isotropic Gaussian. Generative adversarial networks have also been applied to anomaly detection [24, 33]. Pidhorskyi et al. [26] trained AEs with an adversarial loss to detect OoD samples. More recently, introspective variational AEs [8] use the VAE encoder to differentiate between real and reconstructed samples, achieving SOTA image generations and outlier detection performance. Recently, Zhou et al. [39]

investigated the limitations of AEs for OoD. Similarly, we believe that AEs should have two properties: i) *minimality*: the networks should be constrained to only reconstruct ID samples and ii) *sufficiency*: the decoder should have sufficient capacity to reconstruct ID samples with high accuracy. In contrast to [39], where the authors aim at reconstructing only low-dimensional features needed for the classification task, we are interested in reconstructing pseudo-healthy images to enable pixel-wise localization of anomalies.

In this work, we first investigate whether SOTA AEs can learn meaningful representations for anomaly detection. Specifically, we investigate whether AEs can learn the healthy anatomy, i.e., absence of pathology, and generate pseudo-healthy reconstructions of abnormal samples on challenging medical anomaly detection tasks. Our findings are that SOTA AEs either do not efficiently constrain the latent space and allow the reconstruction of anomalous patterns, or that the decoder cannot accurately restore images from their latent representation. The imperfect reconstructions yield high residual errors on normal regions (false positives) that can easily overshadow residuals of interest, i.e., pathology [23]. We then propose *MorphAEus*, novel deformable AEs to learn minimal and sufficient features for anomaly detection and drastically reduce false positives. Figure 1a shows that *MorphAEus* learns the training distribution of healthy chest X-rays and yields pseudo-healthy ID reconstructions even for far OoD samples. This allows to localize pathologies, as seen in Fig. 1b.

Our manuscript advances the understanding of anomaly detection by providing insights into what AEs learn. In summary, our contributions are:

- We broaden the understanding of AEs and highlight their limitations.
- We test whether SOTA AEs can learn the training distribution of the healthy population, accurately reconstruct inputs from their latent representation and reliably detect anomalies.
- As a solution, we propose *MorphAEus*, novel deformable AEs that provide pseudo-healthy reconstructions of abnormal samples and drastically reduce false positives, achieving SOTA unsupervised pathology detection.

2 Background

The widely held popular belief is that AEs can learn the distribution of the training data and identify outliers from inaccurate reconstructions of abnormal samples [31]. This section aims to discuss the common assumptions of unsupervised anomaly detection, specifically for AEs, while also outlining the challenges and desired properties associated with these techniques.

2.1 Unsupervised Anomaly Detection: Assumptions

Let $\mathcal{X} \subset \mathbb{R}^N$ be the data space that describes normal instances for a given task. The manifold hypothesis implies that there exists a low-dimensional manifold $\mathcal{M} \subset \mathbb{R}^D \subset \mathcal{X}$ where all the points $x \in \mathcal{X}$ lie, with $D \ll N$ [9]. For example, a

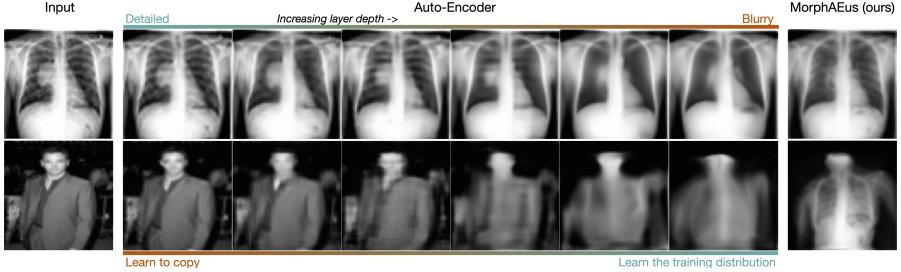


Fig. 2. What do AEs learn? Reconstruction of pathology (top row) and far OoD celebrities (bottom row) with AEs [2] (different depths) and *MorphAEus*. While shallow AEs learn to copy, deep AEs yield blurry reconstructions, rendering AEs unsuitable for anomaly detection. In contrast, *MorphAEus* generates pseudo-healthy ID reconstructions for near (pathology) and far (celebrity) OoD cases.

set of images in pixel space \mathcal{X} could have a compact representation describing features like structure, shape, or orientation in \mathcal{M} .

Given a set of unlabeled data $x_1, \dots, x_n \in \mathcal{X}$ the objective of unsupervised representation learning is to find a function $f : \mathbb{R}^N \rightarrow \mathbb{R}^D$ and its inverse $g : \mathbb{R}^D \rightarrow \mathbb{R}^N$, such that $x \approx g(f(x))$, with the mapping f defining the low-dimensional manifold \mathcal{M} . The core assumption of unsupervised anomaly detection is that once such functions f and g are found, the learned manifold \mathcal{M} would best describe the normal data samples in \mathcal{X} and results in high reconstruction errors for data-points $\bar{x} \notin \mathcal{X}$, that we call anomalous. An anomaly score is therefore usually derived directly from the pixel-wise difference: $s(\bar{x}) = |\bar{x} - g(f(\bar{x}))|$.

The nominal and abnormal distributions are considerably separated from each other when \bar{x} is from a different domain. However, anomalies are often defects in otherwise normal images. In medical imaging, the set \mathcal{X} describes the healthy anatomy and the data set $\bar{\mathcal{X}}$ usually contains images with both healthy and pathological regions. The two distributions usually come from the same domain and might overlap considerably. The core assumption is that only the normal structures can be reconstructed from their latent representation very well, with the pathological regions ideally replaced by healthy structures. Therefore $x \approx g(f(\bar{x})) \in \mathcal{X}$ would represent the healthy synthesis of the abnormal sample \bar{x} and the residual $|\bar{x} - g(f(\bar{x}))|$ would highlight only the abnormal regions.

2.2 Auto-Encoders: Challenges

AEs aim to extract meaningful representations from data, by learning to compress inputs to a lower-dimensional manifold and reconstruct them with minimal error. They use neural networks to learn the functions f and g , often denoted as encoder E_θ with parameters θ and decoder D_ϕ parameterized by a set of parameters ϕ . The embedding $z = E(x|\theta)$ is a projection of the input to a lower-dimensional manifold \mathcal{Z} , also referred to as the bottleneck or latent representation of x . The standard objective of AEs is finding the set of parameters θ

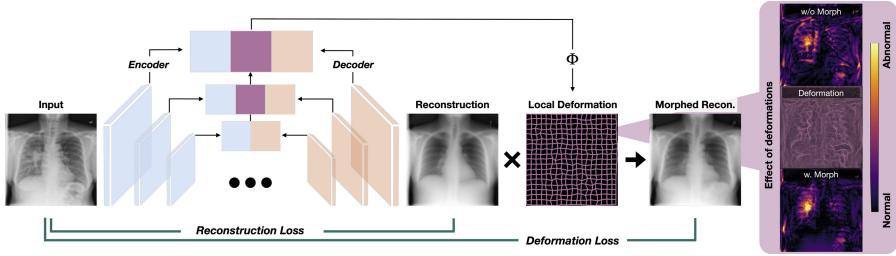


Fig. 3. MorphAEus. Deep perceptual AEs yield reconstructions matching the training distribution. We leverage the top encoding/decoding layers containing spatial information to estimate deep deformation fields. This allows the local morphometric adaptation of the predicted reconstruction to drastically reduce residual errors on normal regions (false positives), as shown on the right.

and ϕ that minimize the residual, with the mean squared error (MSE) being a popular choice for the reconstruction error: $\min_{\theta, \phi} \sum_{i=1}^N \|x_i - D_\phi(E_\theta(x_i))\|^2$.

In the introduction, we presented the desired properties of AEs for outlier detection, namely i) reconstructions should match the training distribution and ii) decoders have sufficient capacity to accurately restore inputs. Figure 2 shows reconstructions of spatial AEs [2] for near OoD samples, i.e., containing pathologies, and far OoD samples using real images of celebrities [19]. AEs with few encoding layers learn to copy and can reconstruct both near- and far OoD. Interestingly, with increasing layer depth, AEs can learn the prior over the training distribution, avoid the reconstruction of pathologies, and project OoD celebrities to the closest chest X-ray counterparts. However, this comes at the cost of losing spatial information and not reconstructing small details, e.g., ribs. The resulting high residual error on healthy tissues (false positives) overshadows the error on pathology [23], rendering standard AEs unsuitable for anomaly detection.

3 MorphAEus: Deformable Auto-encoders

We propose *MorphAEus*, deformable AEs that learn minimal and sufficient features for anomaly detection, see Fig. 3. We use deep perceptual AE to provide pseudo-healthy ID reconstructions and leverage estimated deep deformation fields to drastically reduce the false positives.

Pseudo-healthy Reconstructions. Given a dataset $\mathcal{X} = \{x_1, \dots, x_n\}$ we optimize the encoder and decoder with parameters θ, ϕ to minimize the MSE loss between the input and its reconstruction. For *minimality*, we propose to use deep AEs constrained to reconstruct only ID samples (see Fig. 2), but add a perceptual loss (PL) [14, 37] to encourage reconstructions that are perceptually similar to the training distribution. The reconstruction loss is given by:

$$\mathcal{L}_{Rec}(x|\theta; \phi) = MSE(x, x_{rec}) + \alpha PL(x, x_{rec}), \quad (1)$$

with $x_{rec} = D_\phi(E_\theta(x))$, $PL(x, x_{rec}) = \sum_l (VGG_l(x) - VGG_l(x_{rec}))^2$ with VGG_l being the output of the $l \in \{1, 6, 11, 20\}$ -th layer of a pre-trained VGG-19 encoder. We have empirically found $\alpha = 0.05$ to be a good weight to predict perceptually similar reconstructions without compromising pixel-wise accuracy.

Local Deformation. Imperfect reconstructions yield high residuals on normal regions which might overshadow the residuals errors associated with anomalous regions. Skip connections [30] would allow the network to bypass the learned manifold and copy anomalies at inference. Instead, we propose to align the reconstruction with the input using corresponding encoder and decoder features. We denote these shared parameters as $\theta_S \subset \theta$ and $\phi_S \subset \phi$. Inspired by the advances in image registration [1] and computer vision [4, 36] we estimate dense deformation fields Φ to allow local morphometric adaptations:

$$\mathcal{L}_{Morph}(x|\psi; \theta_S; \phi_S) = 1 - LNCC(x_{morph}, x) + \beta \|\Phi\|^2, \quad (2)$$

where $x_{morph} = x_{rec} \circ \Phi$, ψ are the deformation parameters, LNCC is the local normalized cross correlation, \circ is a spatial transformer and β weights the smoothness constraint on the deformation fields. We opted for the LNCC instead of the MSE to emphasize shape registration and enhance robustness to intensity variations in the inputs and reconstructions. By sharing encoder/decoder parameters, the deformation maps are not only beneficial at inference time, but also guide the training process to learn more accurate features. The full objective is given by the two losses: $\mathcal{L}(x|\theta; \phi; \psi) = \mathcal{L}_{Rec}(x|\theta; \phi) + \mathcal{L}_{Morph}(x|\psi; \theta_S; \phi_S)$. To ensure a good initialization for the deformation estimation, we introduce the deformation loss after 10 epochs. Deformable registration between normal and pathological samples is in itself an active area of research [18, 25]. In particular, the deformation could mask structural abnormalities at inference if not constrained. In our experiments, we linearly increase β from $1e^{-3}$ to 3 to constrain the deformation more as the reconstruction improves (see Appendix for details). Nevertheless, recent advances allow the estimation of the optimal registration parameters automatically [13]. If not specified otherwise, we employ x_{morph} for inference.

4 Pathology Detection on Chest X-rays

In this section, we investigate whether AEs can learn the healthy anatomy, i.e., absence of pathology, and generate pseudo-healthy reconstructions of abnormal chest X-ray images. Pathology detection algorithms are often applied to finding hyper-intense lesions, such as tumors or multiple sclerosis on brain scans. However, it has been shown that thresholding techniques can outperform learning-based methods [22]. In contrast, the detection of pathology on chest radiographs is much more difficult due to the high variability and complexity of nominal features and the diversity and irregularity of abnormalities.

Datasets. We use the Covid-19 radiography database on Kaggle [6, 27]. We used the RSNA dataset [35], which contains 10K CXR images of normal subjects and 6K lung opacity cases. For the detection of Covid-19, we used the Padchest

Table 1. CXR Pathology Detection. Methods marked with ‘*’ achieve improved reconstruction accuracy, but can also reconstruct pathologies which renders them not useful for detecting anomalies, as shown in Fig. 4. In contrast, *MorphAEus* generates only ID reconstructions which enables the localization of pathologies.

Method	Healthy		Pneumonia	Covid-19	Avg.
	SSIM ↑	LPIPS ↓		AUROC ↑	
AE-S [2]	95.9*	1.3*	58.9 ± 0.5	59.8 ± 1.6	59.3
AE-D [2]	75.6	38.3	45.4 ± 0.5	56.1 ± 1.1	50.7
VAE [40]	72.9	38.4	33.5 ± 0.7	43.7 ± 0.7	38.6
β -VAE [12]	63.1	43.9	70.8 ± 0.9	54.5 ± 0.8	62.6
AAE [26]	66.3	15.5	58.4 ± 0.6	59.8 ± 1.2	57.8
SI-VAE [8]	71.3	17.6	49.7 ± 0.7	52.0 ± 0.7	50.9
DAE [16]	95.8*	2.8*	78.0 ± 0.5	76.1 ± 0.7	77.0
MorphAEus (ours)	85.7	9.5	83.6 ± 0.8	86.0 ± 0.7	84.8

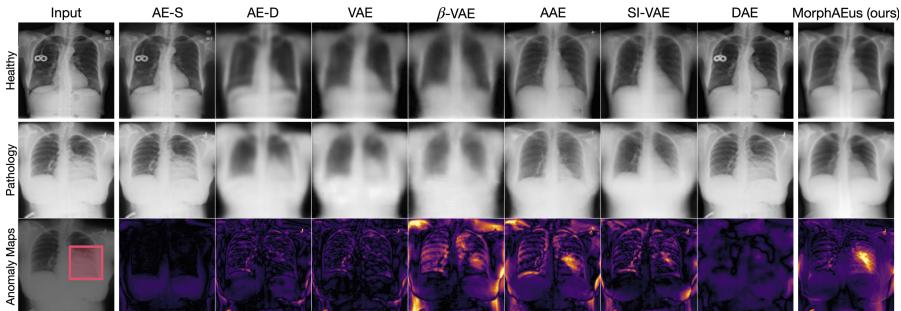


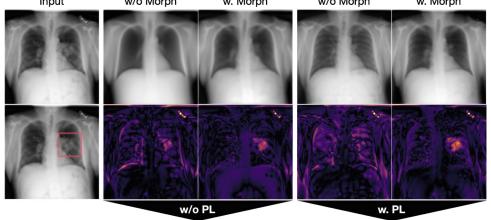
Fig. 4. The first row shows reconstructions of a healthy sample. The next rows show reconstructions of a pathological case and the corresponding anomaly maps. AE-S and DAE fully reconstruct the pathology and thereby fail the task, while the other baselines yield high residual errors on both healthy and pathological regions due to blurry (AE-D, VAE, and β -VAE) or inaccurate (AAE, SI-VAE) reconstructions. Only *MorphAEus* yields accurate pseudo-healthy reconstructions which allows to localize the pathology on the left lung.

dataset [3] containing CXR images manually annotated by trained radiologists. We used 1.3K healthy control images and 2.5K cases of Covid-19.

Results. Of the baselines, only adversarially-trained AEs can reconstruct pseudo-healthy images from abnormal samples, as shown in Fig. 4. However, their imperfect reconstructions overshadow the error on pathology leading to poor anomaly detection results, as reflected in the SSIM and AUROC in Table 1. Spatial AEs and DAEs have a tendency to reproduce the input and produce reconstructions of structures that are not included in the training distribution,

- (a) Morphometric adaptations (M) considerably increase the reconstruction accuracy on healthy (H) and the pathology detection scores (AUROC).

	w/o PL	w. PL	w/o M	w. M	w/o M	w. M
H. (SSIM) \uparrow	80.2	87.9	72.8	85.7		
H. (LPIPS) \downarrow	22.6	13.0	8.3	9.5		
Pathology \uparrow	62.1	82.1	66.8	84.8		



(b) Morphometric adaptations (Morph) reduce false positives and highlight pathological areas.

Fig. 5. Ablation study: Morphometric adaptations considerably improve the pathology detection accuracy with or without perceptual loss (PL).

such as medical devices and pathologies, despite not being trained on OoD data. This can lead to false negatives. DAEs can avoid the reconstruction of some pathologies and achieve good anomaly detection results. However, pathological regions that do not conform to the learned noise model are completely missed, as shown in Fig. 4. The next three methods (AE-D, VAE, and β -VAE) produce blurry reconstructions, as it can be best seen in the LPIPS score. *MorphAEus* is the only method to yield accurate pseudo-healthy reconstructions and effectively remove anomalies, such as pathology or implanted medical devices. This enables to precisely localize pathologies, considerably outperforming the baselines.

Ablation Study: Importance of Morphological Adaptations. We evaluate the effectiveness of individual components of *MorphAEus* in Fig. 5. AEs without perceptual loss tend to not reconstruct small but important features such as ribs and yield false positives on healthy tissue. Interestingly, AEs with perceptual loss achieve more visually appealing reconstructions, but fail at detecting anomalies because the pathological region is overshadowed by false positive residuals on edges and misaligned ribs. Morphometric adaptations guide the networks to learn better representations, reduce the number of false positives, and enable the localization of pathologies. This considerably improves the detection results to 84.8 and 82.1 for AEs with and without perceptual loss, respectively. It is important to note that the deformations are not only beneficial at the time of inference, but also drive the learning process towards better representations. Thereby, the average pathology detection increases from 66.8 to 80.2, even if no adaptations are made during inference, i.e., using x_{rec} instead of x_{morph} for inference (see Appendix for details).

5 Discussion

In this work, we have investigated whether AEs learn meaningful representations to solve pathology detection tasks. We stipulate that AEs should have the desired property of learning the normative distribution (*minimality*) and producing highly accurate reconstructions of ID samples (*sufficiency*). We have

shown that standard, variational, and recent adversarial AEs generally do not satisfy both conditions, nor are they very suitable for pathology detection tasks where the distribution of normal and abnormal instances highly overlap.

In this paper, we introduced *MorphAEus*, a novel deformable AE that demonstrated notable performance improvement in detecting pathology. We believe our method is adaptable to various anomaly types, and we are eager to extend our research to different anatomies and imaging modalities, building upon promising early experiments. However, it is important to address false positive detection, which could be influenced by unlabelled artifacts like medical devices. Our future work aims to conduct a thorough analysis of false positives and explore strategies to mitigate their impact, ultimately enhancing the accuracy.

Although there are obstacles to overcome, AEs remain a viable option for producing easily understandable and interpretable outcomes. Nevertheless, it is crucial to continue improving the quality of the representations to advance unsupervised anomaly detection. Our findings demonstrate that *MorphAEus* is capable of learning superior representations, and can leverage the predicted dense displacement fields to refine its predictions and minimize the occurrence of false positives. This allows for accurate identification and localization of diseases, resulting in SOTA unsupervised pathology detection on chest X-rays.

Acknowledgement. C.I.B. is in part supported by the Helmholtz Association under the joint research school “Munich School for Data Science - MUDS”.

References

1. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: VoxelMorph: a learning framework for deformable medical image registration. *IEEE Trans. Med. Imaging* **38**(8), 1788–1800 (2019)
2. Baur, C., Denner, S., Wiestler, B., Navab, N., Albarqouni, S.: Autoencoders for unsupervised anomaly segmentation in brain MR images: a comparative study. *Med. Image Anal.* **69**, 101952 (2021)
3. Bustos, A., Pertusa, A., Salinas, J.M., de la Iglesia-Vayá, M.: Padchest: a large chest x-ray image dataset with multi-label annotated reports. *Med. Image Anal.* **66**, 101797 (2020)
4. Chen, R., Cong, Y., Dong, J.: Unsupervised dense deformation embedding network for template-free shape correspondence. In: *ICCV*, pp. 8361–8370 (2021)
5. Chen, X., You, S., Tezcan, K.C., Konukoglu, E.: Unsupervised lesion detection via image restoration with a normative prior. *Med. Image Anal.* **64**, 101713 (2020)
6. Chowdhury, M.E.H., et al.: Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access* **8**, 132665–132676 (2020)
7. Ciresan, D., Giusti, A., Gambardella, L., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: *NeurIPS 25* (2012)
8. Daniel, T., Tamar, A.: Soft-introvae: analyzing and improving the introspective variational autoencoder. In: *CVPR*, pp. 4391–4400 (2021)
9. Fefferman, C., Mitter, S., Narayanan, H.: Testing the manifold hypothesis. *J. Am. Math. Soc.* **29**(4), 983–1049 (2016)
10. Golan, I., El-Yaniv, R.: Deep anomaly detection using geometric transformations. In: *NeurIPS vol. 31* (2018)

11. Gong, D., et al.: Memorizing normality to detect anomaly: memory-augmented deep autoencoder for unsupervised anomaly detection. In: ICCV, pp. 1705–1714 (2019)
12. Higgins, I., et al.: beta-VAE: learning basic visual concepts with a constrained variational framework. In: ICLR (2017)
13. Hoopes, A., Hoffmann, M., Fischl, B., Guttag, J., Dalca, A.V.: Hypermorph: amortized hyperparameter learning for image registration. In: IPMI, pp. 3–17 (2021)
14. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: ECCV, pp. 694–711 (2016)
15. Kamnitsas, K., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. **36**, 61–78 (2017)
16. Kascenas, A., Pugeault, N., O’Neil, A.Q.: Denoising autoencoders for unsupervised anomaly detection in brain MRI. In: MIDL (2022)
17. Kingma, D.P., Welling, M.: Auto-encoding variational Bayes. arXiv preprint [arXiv:1312.6114](https://arxiv.org/abs/1312.6114) (2013)
18. Krüger, J., Schultz, S., Handels, H., Ehrhardt, J.: Registration with probabilistic correspondences-accurate and robust registration for pathological and inhomogeneous medical data. CVIU **190**, 102839 (2020)
19. Liu, Z., Luo, P., Wang, X., Tang, X.: Deep learning face attributes in the wild. In: ICCV (2015)
20. Makhzani, A., Frey, B.: k-sparse autoencoders. In: ICLR (2014)
21. Mao, X., Shen, C., Yang, Y.B.: Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In: NeurIPS, vol. 29 (2016)
22. Meissen, F., Kaassis, G., Rueckert, D.: Challenging current semi-supervised anomaly segmentation methods for brain MRI. In: MICCAI brainlesion workshop, pp. 63–74 (2022)
23. Meissen, F., Wiestler, B., Kaassis, G., Rueckert, D.: On the pitfalls of using the residual as anomaly score. In: MIDL (2022)
24. Perera, P., Nallapati, R., Xiang, B.: Ocgan: one-class novelty detection using GANs with constrained latent representations. In: CVPR, pp. 2898–2906 (2019)
25. Periaswamy, S., Farid, H.: Medical image registration with partial data. Med. Image Anal. **10**(3), 452–464 (2006)
26. Pidhorskyi, S., Almohsen, R., Doretto, G.: Generative probabilistic novelty detection with adversarial autoencoders. In: NeurIPS, vol. 31 (2018)
27. Rahman, T., et al.: Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-rays images (2020)
28. Ren, J., et al.: Likelihood ratios for out-of-distribution detection. In: NeurIPS, vol. 32 (2019)
29. Rifai, S., Vincent, P., Muller, X., Glorot, X., Bengio, Y.: Contractive auto-encoders: explicit invariance during feature extraction. In: ICML, pp. 833–840 (2011)
30. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: MICCAI, pp. 234–241 (2015)
31. Ruff, L., Kauffmann, J.R., et al.: A unifying review of deep and shallow anomaly detection. In: Proceeding of IEEE (2021)
32. Schirrmeister, R., Zhou, Y., Ball, T., Zhang, D.: Understanding anomaly detection with deep invertible networks through hierarchies of distributions and features. NeurIPS **33**, 21038–21049 (2020)
33. Schlegl, T., Seeböck, P., Waldstein, S.M., Langs, G., Schmidt-Erfurth, U.: f-AnoGAN: fast unsupervised anomaly detection with generative adversarial networks. Med. Image Anal. **54**, 30–44 (2019)

34. Schlemper, J., Caballero, J., Hajnal, J.V., Price, A.N., Rueckert, D.: A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans. Med. Imaging* **37**(2), 491–503 (2017)
35. Shih, G., et al.: Augmenting the national institutes of health chest radiograph dataset with expert annotations of possible pneumonia. *Radiol. Artif. Intell.* **1**(1), e180041 (2019)
36. Shu, Z., Sahasrabudhe, M., Guler, R.A., Samaras, D., Paragios, N., Kokkinos, I.: Deforming autoencoders: unsupervised disentangling of shape and appearance. In: *ECCV*, pp. 650–665 (2018)
37. Tuluptceva, N., Bakker, B., Fedulova, I., Konushin, A.: Perceptual image anomaly detection. In: *ASPR*, pp. 164–178 (2019)
38. Yoon, S., Noh, Y.K., Park, F.: Autoencoding under normalization constraints. In: *ICML*, pp. 12087–12097 (2021)
39. Zhou, Y.: Rethinking reconstruction autoencoder-based out-of-distribution detection. arXiv preprint [arXiv:2203.02194](https://arxiv.org/abs/2203.02194) (2022)
40. Zimmerer, D., Isensee, F., Petersen, J., Kohl, S., Maier-Hein, K.: Unsupervised anomaly localization using variational auto-encoders. In: *MICCAI* (2019)