Implicit Embeddings via GAN Inversion for High Resolution Chest Radiographs

Tobias Weber[1,2][0000−0002−5430−2595], Michael Ingrisch[2][0000−0003−0268−9078], Bernd Bischl[1][0000−0001−6002−6980], and David Rügamer[1][0000−0002−8772−9202]

[1] Department of Statistics, LMU Munich
{tobias.weber,bernd.bischl,david.ruegamer}@stat.uni-muenchen.de
[2] Department of Radiology, University Hospital, LMU Munich
michael.ingrisch@med.uni-muenchen.de

Abstract. Generative models allow for the creation of highly realistic artificial samples, opening up promising applications in medical imaging. In this work, we propose a multi-stage encoder-based approach to invert the generator of a generative adversarial network (GAN) for high resolution chest radiographs. This gives direct access to its implicitly formed latent space, makes generative models more accessible to researchers, and enables to apply generative techniques to actual patient's images. We investigate various applications for this embedding, including image compression, disentanglement in the encoded dataset, guided image manipulation, and creation of stylized samples. We find that this type of GAN inversion is a promising research direction in the domain of chest radiograph modeling and opens up new ways to combine realistic X-ray sample synthesis with radiological image analysis.

Keywords: generative modeling · latent space disentanglement · representation learning.

1 Introduction

The public release of large datasets for chest radiographs has led to substantial progress in the automated analysis of thoracic X-ray imaging [19, 30, 40]. This availability of large amounts of data facilitates fitting complex generative models with various applications. For example, [15,34] propose an algorithm that creates deceptively real synthetic chest X-ray samples based on generative adversarial

This version of the contribution has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/978-3-031-25046-0_3. Use of this Accepted Version is subject to the publisher's Accepted Manuscript terms of use https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms.

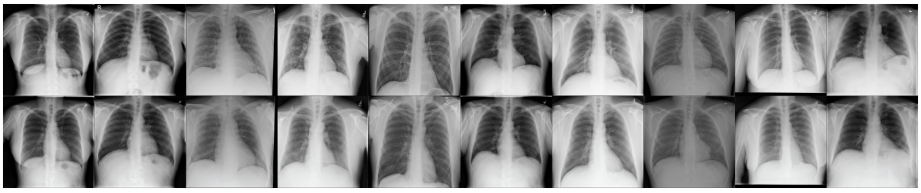


Fig. 1: ChestX-ray14 [40] samples (**top**) and their full resolution reconstructions (**bottom**) from our latent embedding by a GAN-based generator [34].

networks (GANs [13]). In this likelihood-free approach, a generator G is tasked to synthesize fake data from randomly distributed noise $\boldsymbol{z} \in \mathcal{Z}$ in an adversarial setting, while a discriminator network serves as a counterpart that needs to distinguish between the real and fake data.

Previously, GANs were applied in the context of chest radiography as a generative augmentation method to increase the performance of classifiers for underrepresented pathologies [15, 37]. Other applications of generative methods in this domain involve bone suppression [14, 25] and creation of disease saliency maps for abnormal samples [38].

The implicitly formed latent space $\mathcal{Z}$ by GAN-based approaches has been proven to encode disentangled features of the training data with a semantic meaning [12, 16, 22, 35]. A standard GAN model formulation, however, does not allow to access the encoded information.

One remedy for obtaining this latent code is by inverting G and directly projecting data into $\mathcal{Z}$. A possible taxonomy [4, 42] of the GAN inversion topic divides approaches into three groups: 1) Learning-based variants employ an encoder to approximate the embedding [29, 32, 39, 41]. 2) Optimization-based approaches iteratively optimize the latent code $\boldsymbol{z} \in \mathcal{Z}$ for a given target image directly [1, 2, 7, 28]. 3) Hybrid approaches combine the previously mentioned strategies and make use of the encoding as well as the optimization [3]. In the domain of medical imaging, [31] focuses on adjustable mammogram generation via embeddings in the latent space for tumor inpainting. In [11], an inverted GAN assists in converting abdominal computed tomography scans to magnetic resonance images and vice versa. In contrast, [33] apply the reverted generation process as a proxy to increase output probabilities in various target classes for explaining deep black-box classifiers.

Our Contribution In this paper, we propose a novel multi-stage hybrid approach with bootstrapped pre-training for aligning the encoder directly to the distribution of the generator to map thoracic X-ray images into $\mathcal{Z}$. Moreover, we elaborate on practical applications and implications of this implicitly created embedding space in the domain of chest radiography on a full dataset scale. This includes the aspect of image compression (cf. Figure 1), disentangled encoding in the latent space, and the ability to perform image manipulations beyond synthetic sample generation. Our method allows us to, e.g., model the course of

thorax or lung diseases on the actual radiographs on real-world patient data, or create stylized or similar samples for a given target chest X-ray image using the reverse encoding.

2 Methods

In GANs, the generator defines a mapping $G : \mathcal{Z} \to \mathcal{X}$, where $\mathcal{X}$ is an arbitrary data space and $\mathcal{Z}$ forms the latent space. In contrast to most $\mathcal{X}$, the latent space is considered as smooth and implicitly encodes rich disentangled semantic features in a low-dimensional manifold with $\dim(\mathcal{Z}) \ll \dim(\mathcal{X})$. However, due to the nature of GAN training and the non-invertibility of neural networks, the mapping from $\mathcal{X}$ to $\mathcal{Z}$ is unknown. This problem can also be portrayed as

$$\boldsymbol{z}^* = \arg\min_{\boldsymbol{z} \in \mathcal{Z}} \mathcal{L}(G(\boldsymbol{z}), \boldsymbol{x}) \,, \tag{1}$$

where $\boldsymbol{z}^*$ minimizes a pre-defined criterion $\mathcal{L} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such as the mean squared error (MSE) for $\boldsymbol{z} \in \mathcal{Z}$ and $\boldsymbol{x} \in \mathcal{X}$. In this work, we employ a three-stage hybrid approach to estimate the ideal latent code $\boldsymbol{z}^*$ (cf. Figure 2).

Stage 1: Bootstrapped Training As a first step, we use an encoder $E : \mathcal{X} \to \mathcal{Z}$ to approximate the inversion of G. To account for the limited availability of training data, we propose to exploit the generation capabilities of G by sampling $\boldsymbol{z}$ from a pre-defined distribution $p_{\boldsymbol{z}}$ (e.g., standard normal distribution) and producing new $\boldsymbol{x} = G(\boldsymbol{z})$. This composition allows for training E on the latent code directly and on a theoretically infinite amount of data. [5] also show that $\boldsymbol{z}$ already represents compositional properties of $\boldsymbol{x}$ and a strong image prior is formed by the regressor. We define the objective for one weight update of E with a batch

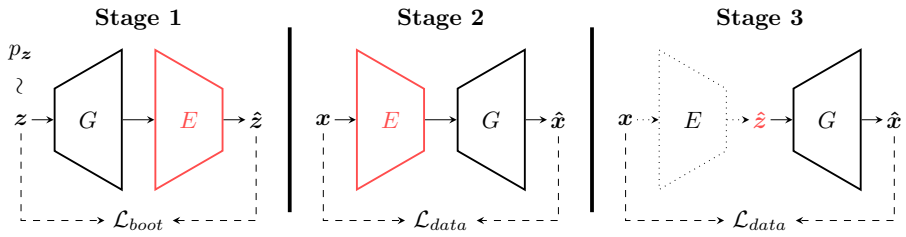


Fig. 2: Overview of our three-stage hybrid inversion approach. The red color symbolizes the part of the architecture with trainable parameters. The pre-training phase (Stage 1, **left**) learns a latent code representation of artificial samples. In Stage 2 (**middle**) the encoder E is finetuned with real data samples $\boldsymbol{x}$. Lastly, each estimated $\hat{\boldsymbol{z}} = E(\boldsymbol{x})$ is iteratively refined with frozen network components (**right**).

size of B as

$$\mathcal{L}_{boot} = \frac{1}{B}\sum_{i=1}^{B} |\boldsymbol{z}_i - E(G(\boldsymbol{z}_i))| \tag{2}$$

with $|\cdot|$ being the l_1 norm and $\boldsymbol{z}_i \sim p_{\boldsymbol{z}}$. We hypothesize that this bootstrapped form of pre-training leads to a better generalization of E as the generator is also able to generate meaningful edge- and out-of-distribution samples.

Stage 2: Dataset Training In the second stage of our approach, E is finetuned on real data observations $\boldsymbol{x} \in \mathcal{X}$ using the loss $\mathcal{L}_{data}$, a criterion for the distance of $\boldsymbol{x}$ and $\hat{\boldsymbol{x}} = G(E(\boldsymbol{x}))$. More specifically, since $\mathcal{X}$ is assumed to reside in the image domain, we employ multiple losses as in [1, 39] that account for various levels of similarity: $\mathcal{L}_{MOCO}$ minimizes the angle of feature embeddings produced by a contrastive network C for true and predicted observations $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$. It is defined as

$$\mathcal{L}_{MOCO} = 1 - C(\boldsymbol{x})^\intercal C(\hat{\boldsymbol{x}})\,. \tag{3}$$

Following [39], C is a ResNet-50 [18] trained on Image-net [8] with MOCOv2 [6, 17]. To enforce perceptual similarity, $\mathcal{L}_{LPIPS}$ [43] is chosen, which compares weight-scaled activations of $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ in hidden layers of a pre-trained VGG16 [36] network, i.e. it aligns deep features of original and reconstruction. Lastly, $\mathcal{L}_{MSE}$ serves as a measure for the pixel-wise distance. We combine these losses into

$$\mathcal{L}_{data} = \lambda_{MSE}\mathcal{L}_{MSE} + \lambda_{LPIPS}\mathcal{L}_{LPIPS} + \lambda_{MOCO}\mathcal{L}_{MOCO}\,, \tag{4}$$

where the regularization parameters $\lambda_\cdot \in \mathbb{R}^+$ are used to weigh the different losses.

Stage 3: Iterative Optimization The area of interest in medical imaging is often only a small part of the full image. To further enhance the quality of the inversion, we utilize gradient-based per sample optimization of (1) to obtain a latent code close to $\boldsymbol{z}^*$ for selected images. Here, $\boldsymbol{z} = E(\boldsymbol{x})$ and Eq. (1) is used as objective function with $\mathcal{L}_{data}$ as criterion. Note that this optimization only involves $\boldsymbol{z}$ as trainable parameter.

3 Experiments

The basis for our experiments is the ChestX-ray14 dataset [40] with over 100k images in 1024×1024 pixel resolution of 30k patients and 14 labeled pathologies. We build on the work of [34], which provides a progressive growing generator [20] for the generation of synthetic chest radiographs as G with $\dim(\mathcal{Z}) = 512$ and $p_{\boldsymbol{z}} = \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. E utilizes ConvNext-small [26] as a backbone, where we replace the last layer with a fully connected layer, whose output dimension matches the dimension of the latent space. Stage 1 is trained with a batch size of 64 and 500k batch updates. E is then finetuned for 15 epochs on the dataset in stage 2. Both stages use the Adam [24] optimizer with learning rates of 5e−5 and 1e−5, respectively. Iterative optimization is done with 3000 iterations.

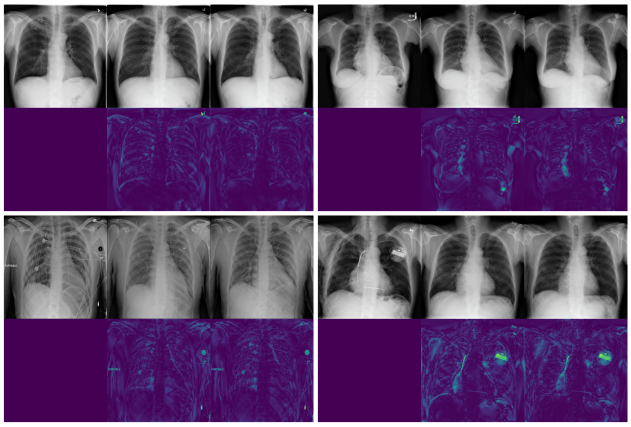


Fig. 3: Reconstruction capabilities of E. The four examples show the original image (**left** column) as well as the reconstruction from $\boldsymbol{z}$ estimated by E (**central** column) and after iterative optimization (**right** column). In each case, the **bottom row** visualizes the residual from the original with purple being a zero value and brighter values correspond to larger errors.

3.1 Image Compression and Quality of Reconstruction

One of the main criteria for a successful inversion is the reconstruction quality from the obtained $\boldsymbol{z}$. In the best-case scenario, E and G enforce a cycle consistency [9, 44], i.e., a perfect inversion. In Figure 3 different examples from the test set are depicted. The inversion process captures the majority of key image features, while the iterative optimization results in smaller residuals compared to the single encoder pass. Quantitatively, this evaluates to an increase of the structural similarity index measure from 0.813 to 0.825 and a peak signal-to-noise ratio from 22.28 to 23.95 on the test dataset. The reconstruction is bounded by the capacity of G and as stated in [34], the synthetic samples lack some details of image annotations and external medical devices like chest tubes or pacemakers. We observe that these artifacts are simply missing in the reconstructed image. In this application, GAN inversion can be understood as a special form of compression as the high-resolution input is condensed into a semantic 512-dimensional vector, which is $\approx 0.05\%$ of its original size. While compression requires additional expert knowledge to assure that all relevant information of the original images is retained in $\mathcal{Z}$, the ad-hoc identification of spurious features is another application of the inversion method and allows medical laypersons to quickly identify abnormalities in the data.

3.2 Disentanglement in Latent Space

Mapping a complex and sparse $\mathcal{X}$ to the latent $\mathcal{Z}$ results in a smooth and highly semantic representation [35], which allows for further disentanglement analysis

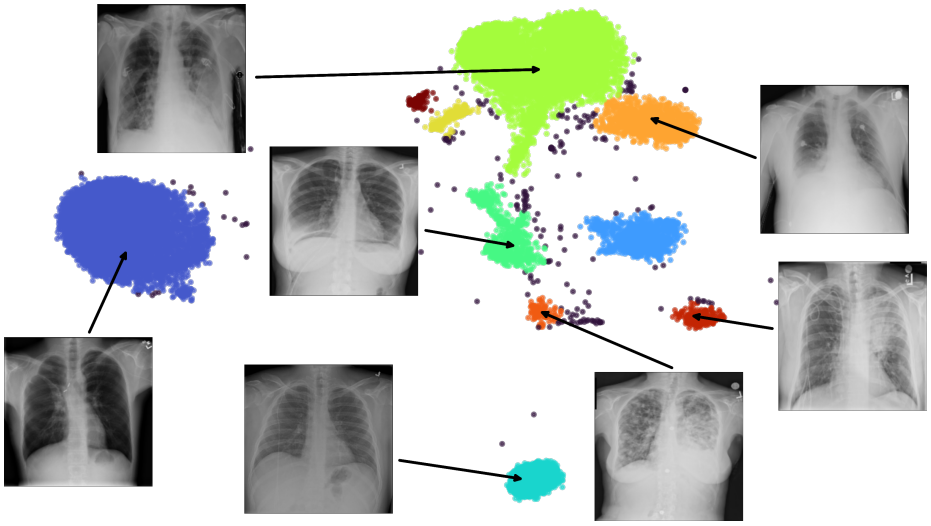
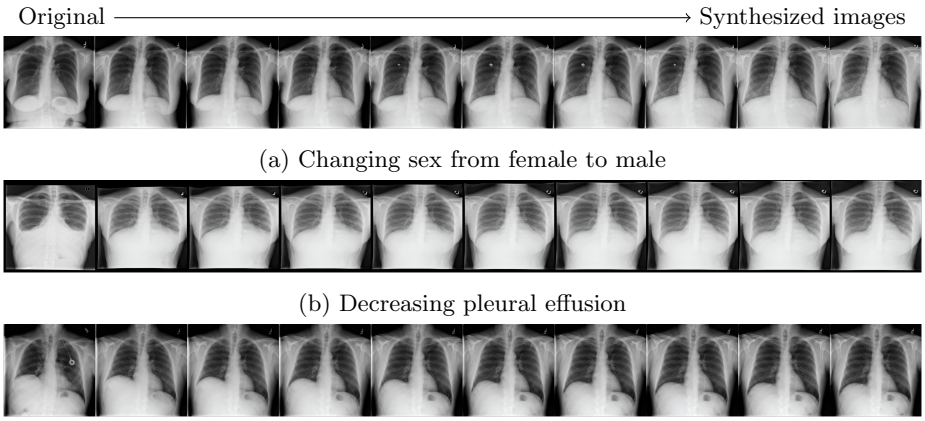


Fig. 4: UMAP embedding (point clouds) of the test set's latent code. Selected samples show key characteristics of various clusters.

of the radiograph's features. In particular, grouping distinctive features becomes feasible in this low-dimensional space. We use the non-linear UMAP [27] to further condense the latent code and apply density-based clustering (DBSCAN, [10]) to isolate groups, resulting in Figure 4. Visual exploration suggests that most variation of the data stems from the sex of patients as well as the image contrast, signal intensity, and the patient's posture. These attributes mostly form coherent clusters. An analysis of the reduced embedding based on certain pathologies proves difficult, as these are often high-frequency features with low importance for the image reconstruction and account only for minimal variation in $\mathcal{Z}$.

3.3 Guided Image Manipulation

Disentangled semantics in $\mathcal{Z}$ allow manipulating explicit features of an individual image $\boldsymbol{x}$. We investigate the technique of InterFaceGAN [35] for facial editing applied to our use case. Based on a pre-specified target attribute, a support vector machine (SVM) with a linear kernel is trained to separate the latent codes of the dataset, which were obtained by E. The normal vector $\mathbf{n}$ of the resulting hyperplane then serves as interpolation axis and allows to generate new latent codes $\mathbf{z}_{\text{new}}$ along the direction of most class diversity: $\boldsymbol{z}_{\text{new}} = \boldsymbol{z} + \alpha\mathbf{n}$, where $\alpha \in \mathbb{R}$ is a scaling factor. By increasing α, the chosen target attribute in the generated image from $\boldsymbol{z}_{\text{new}}$ will be more pronounced and vice versa, while other features of the image remain unchanged. Figure 5 exemplarily explores the manipulation of a patient's sex, pleural effusion, and atelectasis. We observe that conditions with a large impact on the full image, such as atelectasis, pleural



(a) Changing sex from female to male

(b) Decreasing pleural effusion

(c) Decreasing atelectasis

Fig. 5: Manipulation of given radiographs based on sex (**top**), pleural effusion (**middle**) and atelectasis (**bottom**). Starting from the original image (**left**), an interpolation along $\mathbf{n}$ shows the desired change.

effusion, or cardiomegaly, are suitable for this type of manipulation. Diseases on smaller scales such as lung nodules, on the other hand, are not well captured by the latent embedding and are thus more challenging to manipulate.

3.4 Proximity sampling

Aside from guided image manipulation, the latent embedding of a radiograph can also be used to synthesize new samples with the style and core features of its original image. This can be achieved by sampling new latent vectors in the proximity of the targeted $\boldsymbol{z}$ and reconstructing an image from the sample with G as depicted in Figure 6. Random samples in the close neighborhood of the original image show minor variations of key features. The further the distance, the more variation is observed in the generated output image.

4 Outlook and Conclusion

While producing highly realistic samples, resulting images of our inversion process miss medical devices and often fine-grained details that are crucial for diagnosing pathologies. A potential solution to this could be a different backbone architecture such as current state-of-the-art style-based architectures [21–23] instead of relying on PGAN [20] as in [34]. These GAN variants do not rely on a parametric space assumption, but learn a representation $\mathcal{W}$ of $\mathcal{Z}$ using a non-linear map and allow the extension to $\mathcal{W}+$ with layer-specific latent codes in the subsequent inversion task. This results in additional out-of-distribution generalizability for better reconstruction quality and style mixing [42]. Utilizing

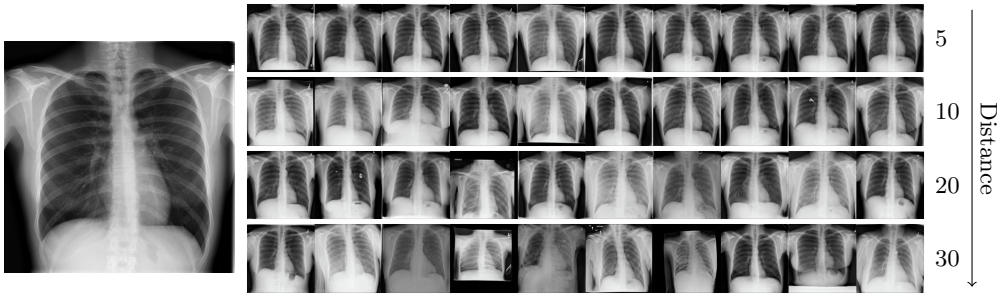


Fig. 6: Generating various samples in the neighborhood of one example scan (**left**). The **right** grid shows samples with a growing Euclidean distance to the original.

$\mathcal{W}$ contributes to enhanced linear separability with respect to certain key features and promotes disentanglement [22], which in turn could be decisive for our classifier-based image traversing on pathologies. It could also be beneficial for future research to enlarge the training data by, e.g., incorporating the CheXpert database [19].

In this paper, we examined the capability and opportunities of GAN inversion in the context of chest radiographs. We employed a multi-stage hybrid procedure, which utilizes both an encoder and an iterative optimization to map high-resolution images to latent code. Furthermore, we have shown that GAN inversion can be used to explore the implicit latent representation of chest X-ray images, and have demonstrated applications such as data compression, disentanglement in the latent space, guided image manipulation, and synthesizing of stylized samples.

Acknowledgments

This work has been funded by the German Federal Ministry of Education and Research and the Bavarian State Ministry for Science and the Arts. The authors of this work take full responsibility for its content. The authors gratefully acknowledge LMU Klinikum for providing computing resources on their Clinical Open Research Engine (CORE). We thank the anonymous reviewers for their constructive comments, which helped us to improve the manuscript.

References

1. Abdal, R., Qin, Y., Wonka, P.: Image2StyleGAN: How to Embed Images Into the StyleGAN Latent Space? In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4432–4441 (2019)

2. Abdal, R., Qin, Y., Wonka, P.: Image2StyleGAN++: How to Edit the Embedded Images? In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8296–8305 (2020)
3. Bau, D., Zhu, J.Y., Wulff, J., Peebles, W., Strobelt, H., Zhou, B., Torralba, A.: Seeing What a GAN Cannot Generate. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4502–4511 (2019)
4. Bermano, A.H., Gal, R., Alaluf, Y., Mokady, R., Nitzan, Y., Tov, O., Patashnik, O., Cohen-Or, D.: State-of-the-Art in the Architecture, Methods and Applications of StyleGAN. In: Computer Graphics Forum. vol. 41, pp. 591–611. Wiley Online Library (2022)
5. Chai, L., Wulff, J., Isola, P.: Using latent space regression to analyze and leverage compositionality in GANs. In: 9th International Conference on Learning Representations, ICLR (2021)
6. Chen, X., Fan, H., Girshick, R., He, K.: Improved Baselines with Momentum Contrastive Learning. arXiv:2003.04297 [cs] (2020)
7. Creswell, A., Bharath, A.A.: Inverting The Generator Of A Generative Adversarial Network. IEEE transactions on neural networks and learning systems **30**(7), 1967–1974 (2018)
8. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
9. Donahue, J., Krähenbühl, P., Darrell, T.: Adversarial Feature Learning. In: 5th International Conference on Learning Representations, ICLR (2017)
10. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. pp. 226–231. KDD'96, AAAI Press, Portland, Oregon (1996)
11. Fetty, L., Bylund, M., Kuess, P., Heilemann, G., Nyholm, T., Georg, D., Löfstedt, T.: Latent space manipulation for high-resolution medical image synthesis via the StyleGAN. Zeitschrift für Medizinische Physik **30**(4), 305–314 (2020)
12. Goetschalckx, L., Andonian, A., Oliva, A., Isola, P.: GANalyze: Toward Visual Definitions of Cognitive Image Properties. In: Proceedings of the ieee/cvf international conference on computer vision. pp. 5744–5753 (2019)
13. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative Adversarial Networks. Advances in neural information processing systems **27** (2014)
14. Han, L., Lyu, Y., Peng, C., Zhou, S.K.: Gan-based disentanglement learning for chest x-ray rib suppression. Medical Image Anal. **77** (2022)
15. Han, T., Nebelung, S., Haarburger, C., Horst, N., Reinartz, S., Merhof, D., Kiessling, F., Schulz, V., Truhn, D.: Breaking medical data sharing boundaries by using synthesized radiographs. Science advances **6**(49) (2020)
16. Härkönen, E., Hertzmann, A., Lehtinen, J., Paris, S.: GANSpace: Discovering Interpretable GAN Controls. Advances in Neural Information Processing Systems **33**, 9841–9850 (2020)
17. He, K., Fan, H., Wu, Y., Xie, S., Girshick, R.: Momentum Contrast for Unsupervised Visual Representation Learning. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 9729–9738 (2020)
18. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)

19. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R., Shpanskaya, K., et al.: CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. In: Proceedings of the AAAI conference on artificial intelligence. vol. 33, pp. 590–597 (2019)
20. Karras, T., Aila, T., Laine, S., Lehtinen, J.: Progressive Growing of GANs for Improved Quality, Stability, and Variation. In: 6th International Conference on Learning Representations, ICLR (2018)
21. Karras, T., Aittala, M., Laine, S., Härkönen, E., Hellsten, J., Lehtinen, J., Aila, T.: Alias-Free Generative Adversarial Networks. Advances in Neural Information Processing Systems **34** (2021)
22. Karras, T., Laine, S., Aila, T.: A Style-Based Generator Architecture for Generative Adversarial Networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4401–4410 (2019)
23. Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T.: Analyzing and Improving the Image Quality of StyleGAN. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8110–8119 (2020)
24. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization. In: 3rd International Conference on Learning Representations, ICLR (2017)
25. Li, Z., Li, H., Han, H., Shi, G., Wang, J., Zhou, S.K.: Encoding CT anatomy knowledge for unpaired chest x-ray image decomposition. In: Medical Image Computing and Computer Assisted Intervention - MICCAI 2019 - 22nd International Conference. Lecture Notes in Computer Science, vol. 11769, pp. 275–283. Springer (2019)
26. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A ConvNet for the 2020s. arXiv:2201.03545 [cs] (2022)
27. McInnes, L., Healy, J., Saul, N., Großberger, L.: UMAP: uniform manifold approximation and projection. J. Open Source Softw. **3**(29), 861 (2018)
28. Menon, S., Damian, A., Hu, S., Ravi, N., Rudin, C.: PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 2437–2445 (2020)
29. Nitzan, Y., Bermano, A., Li, Y., Cohen-Or, D.: Face identity disentanglement via latent space mapping. ACM Trans. Graph. **39**(6), 225:1–225:14 (2020)
30. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M.P., Ng, A.Y.: CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv:1711.05225 [cs, stat] (2017)
31. Ren, Z., Yu, S.X., Whitney, D.: Controllable Medical Image Generation via GAN. Journal of Perceptual Imaging **5**(0), 000502–1–000502–15 (2022)
32. Richardson, E., Alaluf, Y., Patashnik, O., Nitzan, Y., Azar, Y., Shapiro, S., Cohen-Or, D.: Encoding in Style: a StyleGAN Encoder for Image-to-Image Translation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2287–2296 (2021)
33. Schutte, K., Moindrot, O., Hérent, P., Schiratti, J.B., Jégou, S.: Using StyleGAN for Visual Interpretability of Deep Learning Models on Medical Images. arXiv:2101.07563 [cs, eess] (2021)
34. Segal, B., Rubin, D.M., Rubin, G., Pantanowitz, A.: Evaluating the Clinical Realism of Synthetic Chest X-Rays Generated Using Progressively Growing GANs. SN Computer Science **2**(4), 1–17 (2021)

35. Shen, Y., Gu, J., Tang, X., Zhou, B.: Interpreting the Latent Space of GANs for Semantic Face Editing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9243–9252 (2020)
36. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations, ICLR (2015)
37. Sundaram, S., Hulkund, N.: Gan-based data augmentation for chest x-ray classification. arXiv:2107.02970 (2021)
38. Tang, Y., Tang, Y., Zhu, Y., Xiao, J., Summers, R.M.: A disentangled generative model for disease decomposition in chest x-rays via normal image synthesis. Medical Image Anal. **67** (2021)
39. Tov, O., Alaluf, Y., Nitzan, Y., Patashnik, O., Cohen-Or, D.: Designing an Encoder for StyleGAN Image Manipulation. ACM Transactions on Graphics (TOG) **40**(4), 1–14 (2021)
40. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2097–2106 (2017)
41. Wei, T., Chen, D., Zhou, W., Liao, J., Zhang, W., Yuan, L., Hua, G., Yu, N.: E2Style: Improve the Efficiency and Effectiveness of StyleGAN Inversion. IEEE Transactions on Image Processing **31**, 3267–3280 (2022)
42. Xia, W., Zhang, Y., Yang, Y., Xue, J.H., Zhou, B., Yang, M.H.: GAN Inversion: A Survey. arXiv:2101.05278 [cs] (2022)
43. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 586–595 (2018)
44. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2223–2232 (2017)