SALAD: Self-supervised Aggregation Learning for Anomaly Detection on X-Rays

Behzad Bozorgtabar[1,2,3](✉), Dwarikanath Mahapatra[4], Guillaume Vray[1], and Jean-Philippe Thiran[1,2,3]

[1] Signal Processing Laboratory 5 EPFL, Lausanne, Switzerland
{behzad.bozorgtabar,guillaume.vray,jean-philipe.thiran}@epfl.ch
[2] Department of Radiology, Lausanne University Hospital, Lausanne, Switzerland
[3] Center of Biomedical Imaging, Lausanne, Switzerland
[4] Inception Institute of Artificial Intelligence, Abu Dhabi, UAE

Abstract. Deep anomaly detection models using a supervised mode of learning usually work under a closed set assumption and suffer from overfitting to previously seen rare anomalies at training, which hinders their applicability in a real scenario. In addition, obtaining annotations for X-rays is very time consuming and requires extensive training of radiologists. Hence, training anomaly detection in a fully unsupervised or self-supervised fashion would be advantageous, allowing a significant reduction of time spent on the report by radiologists. In this paper, we present SALAD, an end-to-end deep self-supervised methodology for anomaly detection on X-Ray images. The proposed method is based on an optimization strategy in which a deep neural network is encouraged to represent prototypical local patterns of the normal data in the embedding space. During training, we record the prototypical patterns of normal training samples via a memory bank. Our anomaly score is then derived by measuring similarity to a weighted combination of normal prototypical patterns within a memory bank without using any anomalous patterns. We present extensive experiments on the challenging NIH Chest X-rays and MURA dataset, which indicate that our algorithm improves state-of-the-art methods by a wide margin.

Keywords: Anomaly detection · X-rays · Self-supervised learning · Deep similarity metric

1 Introduction

Currently, supervised based deep learning approaches are ubiquitous and achieve promising results for abnormality detection in X-ray images [21]. However, many

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-59710-8_46) contains supplementary material, which is available to authorized users.

A. L. Martel et al. (Eds.): MICCAI 2020, LNCS 12261, pp. 468–478, 2020.
https://doi.org/10.1007/978-3-030-59710-8_46

real-world datasets of radiographs often have long-tailed label distributions. On these datasets, deep neural networks have been found to perform poorly on rare classes of anomalies. This particularly has a pernicious effect on the deployed model if, at test time, we place more emphasis on minority classes of abnormal X-ray images. For example, in detecting rare lung opacities, e.g., such as pneumonia in chest X-rays (CXR), normal X-rays are much easier to acquire. Besides, examining radiographs and reporting work for the signs of abnormalities are very time consuming and require qualified radiologists.

Anomaly detection based methods [1,2,8,13] can be significantly useful in large-scale disease screening and spotting candidate regions for anomalies. Classical anomaly detection (AD) methods such as One-Class SVM (OC-SVM) [19], Local Outlier Factor [3], or Isolation Forest [12] often fail to be effective on high-dimensional data or be scaled to large datasets. To alleviate this concern, state-of-the-art methods like Deep SVDD [16] and Deep SAD [17] consider learning deep CNN features as an alternative to classical one-class anomaly detection. However, the former suffers from the well-known problem of mode collapse, while the latter tends to ignore the underlying structure of the images as the pre-trained weights from autoencoder are sensitive to biased low-level features. Reconstruction-based methods [8,23,25] use well-established convolutional autoencoders to compress and reconstruct single-class normal samples, but autoencoders can sometimes reconstruct abnormal samples well, yielding miss detection of anomalies at test time. New methods built upon generative adversarial networks (GANs) [5,18,20] have shown promising anomaly detection performance by using GANs' ability to learn a manifold of normal samples. However, generated samples by GANs do not always lie at the boundary of real data distribution, which is necessary to distinguish normal images from abnormal ones. Recently, self-supervised methods [4,7,10,24] have been proposed to use unlabeled data in a task-agnostic way for extracting generalizable features, where the dataset can be labeled by exploiting the relations between different input samples, rather than requiring external labels. For example, self-supervised deep methods [6,7] proposed to train a classifier for which a self-labeled multi-class dataset is created by applying a set of geometric transformations to the images. However, these methods are domain-specific and cannot generalize over other data types.

Contribution. In this paper, we propose SALAD, short for **S**elf-supervised **A**ggregation **L**earning for **A**nomaly **D**etection on X-rays, a new training scheme that derives an aggregation learning from measuring the similarity between the estimated features of normal samples, to improve clustering and form prototypical patterns. We present a principled formulation to bypass tedious annotations and remove potential bias introduced by training. We show our method's superiority to existing anomaly detection methods on X-ray datasets. We also highlight the limitations of the current state-of-the-art methods.

2 Method

The merit of our approach is self-supervised representation learning, where the feature representation of each X-ray image is pushed closer to its similar neighbors, forming well-clustered features (prototypical patterns) in the latent space. This intuition is illustrated in Fig. 1. To do so, we propose to minimize the *entropy* of each sample feature point's similarity distribution to other nearby samples. Learning feature similarity would require obtaining image embedding in the entire dataset. To avoid this, we use a memory bank [22] to record and use the features. In every iteration, the memory bank is updated with the mini-batch features. The clustering objective will help us identify abnormal samples if they have different characteristics compared to normal prototypical patterns.

The backbone of our anomaly detection model is based on deep auto-encoder, where the encoder $f_{\theta_{enc}} : \mathcal{X} \rightarrow \mathcal{Z}$ is a convolutional neural network that represents input images $\{x_i \in \mathcal{X}\}_{i=1}^{N}$ in an informative latent domain $\mathcal{Z}$. The encoded representation performs as a query to compare with the relevant items in the feature memory bank. The decoder $g_{\theta_{dec}} : \mathcal{Z} \rightarrow \mathcal{X}$ is an up-sampling convolutional neural network that reconstructs the samples given their latent representations.

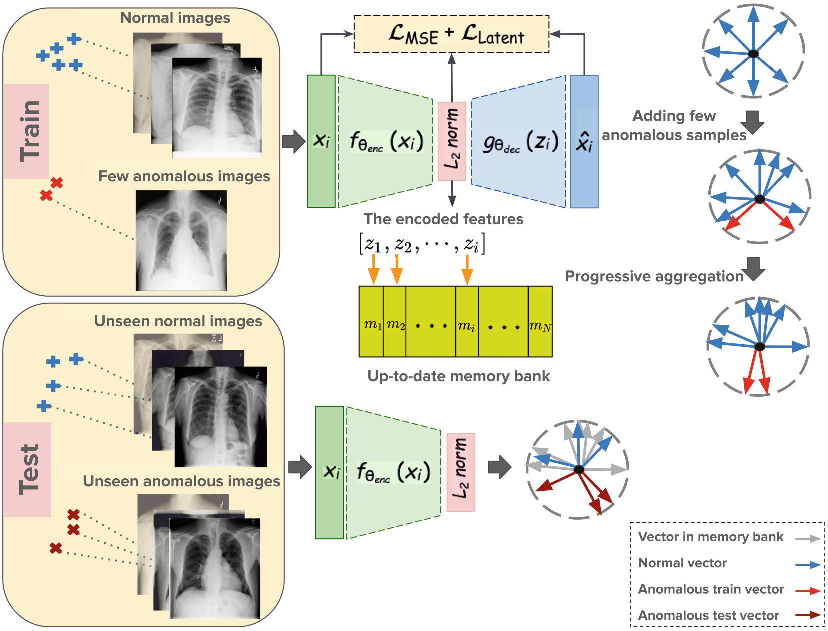


Fig. 1. The SALAD pipeline. The training process starts with forming the prototypical normal patterns (top). At test time, we measure the similarity between the test sample and normal patterns recorded in a memory bank (bottom).

Pre-training. For initialization, we establish an autoencoder pre-training routine using the image reconstruction loss (mean squared error), i.e. $\mathcal{L}_{\text{mse}} = \min\limits_{\theta_{enc},\theta_{dec}} \|x - g_{\theta_{dec}} \circ f_{\theta_{enc}}(x)\|_2^2$. In addition, we impose a constraint on the lower-dimensional representation of the data in which features of the same X-ray image under random data augmentations are invariant, while the features of different images are scattered. To do so, we jointly optimize the training of network with reconstruction loss and a sample specific loss $\mathcal{L}_{\text{ss}}$ [22] to enforce a unique representation for each image:

$$\min_{\theta_{enc}} \mathcal{L}_{\text{ss}} = -\sum_{i \in \mathcal{B}_{\text{spl}}} \log\Big(\sum_{j \in AUG_i} p_{i,j}\Big) \quad \text{s.t.} \quad p_{i,j} = \frac{\exp\left(z_j^T z_i/\tau\right)}{\sum_{k=1}^{N} \exp\left(z_k^T z_i/\tau\right)} \tag{1}$$

where $\tau \in (0,1]$ denotes a fixed temperature hyperparameter. $\mathcal{B}_{\text{spl}}$ denotes the set of samples in the mini-batch. z_i is the feature representation and AUG_i denotes the set of randomly augmented versions of the image x_i.

Training. The learned feature representation at the pre-training stage may not preserve the similarity of different images. Therefore, we add the aggregation loss $\mathcal{L}_{\text{agg}}$ (Eq. 2) to enforce consistency between samples lying in a neighborhood in latent space. We define aggregation loss as the (negative) log-likelihood that a specific sample will be identified as a member of the set of adjacent samples sharing the same prototypical pattern. This is achieved by the entropy measurement of the probability vector in Eq. 1. The more similar the samples are, the less relative entropy they have. We progressively increase entropy to consider larger prototypical neighborhood for the samples and form clusters (see Fig. 2). Finally, the proposed loss $\mathcal{L}_{\text{salad}}$ (Eq. 3) joins all training losses:

$$\min_{\theta_{enc}} \mathcal{L}_{\text{agg}} = -\sum_{i \in \mathcal{B}_{\text{ps}}} \log\Big(\sum_{j \in \mathcal{N}_k(z_i)} p_{i,j}\Big) \tag{2}$$

$$\min_{\theta_{enc},\theta_{dec}} \mathcal{L}_{\text{salad}} = \min_{\theta_{enc},\theta_{dec}} \mathcal{L}_{\text{mse}} + \lambda \min_{\theta_{enc}} \overbrace{(\mathcal{L}_{\text{ss}} + \mathcal{L}_{\text{agg}})}^{\mathcal{L}_{\text{latent}}} \tag{3}$$

where $\mathcal{N}_k(z_i)$ denotes the top-k neighbours determined by the lowest cosine distance with respect to the embedding vector z_i. λ is a hyperparameter to scale the losses ($\mathcal{L}_{\text{latent}}$) used in the latent space and $\mathcal{B}_{\text{ps}}$ denotes the set of prototypical samples in a mini-batch.

Memory Bank. Similarly to [22,24], we first initialize the memory bank with random unit vectors and then update its values m_i using a weighted moving average scheme $m_i \leftarrow (1-t)\, m_i + t z_i$ considering the up-to-date features z_i, where t is the fixed hyperparameter.

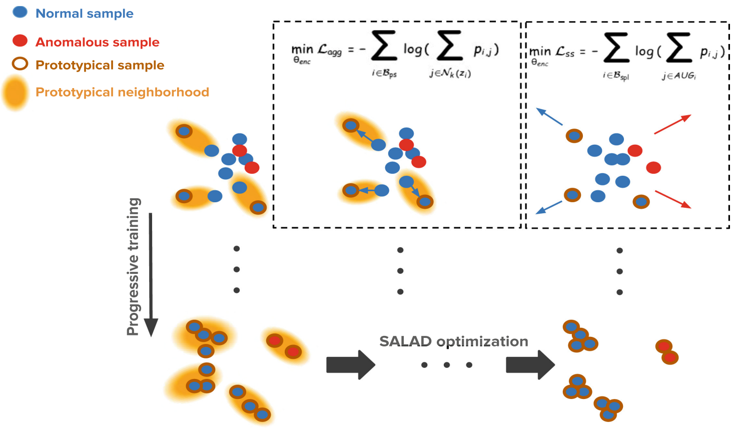

Normal sample
Anomalous sample
Prototypical sample
Prototypical neighborhood
Progressive training
$\min_{\theta_{enc}} \mathcal{L}_{agg} = -\sum_{i\in\mathcal{B}_{ps}} \log(\sum_{j\in\mathcal{N}_k(z_i)} p_{i,j})$
$\min_{\theta_{enc}} \mathcal{L}_{ss} = -\sum_{i\in\mathcal{B}_{spl}} \log(\sum_{j\in AUG_i} p_{i,j})$
SALAD optimization

Fig. 2. An overview of a proposed progressive training strategy. We gradually increase sample neighborhoods to form prototypical patterns.

Inference. In the testing phase, an X-ray image is passed through the trained encoder and its representation is compared with the most relevant normal patterns in the memory bank for computing an anomaly score. Motivated by the weighted k-nearest neighbors (kNN), each vote w_i is obtained from the top K nearest feature vectors in the memory bank. An anomaly score $\mathcal{A}(\cdot)$ is calculated by:

$$\mathcal{A}(x_i) = \frac{1}{K}\sum_{k=1}^{K} w_{i,k} \quad \text{s.t.} \quad w_{i,k} = \frac{\arccos\left(d\left(z_i, m_k\right)\right)}{\sum_{j=1}^{N} \arccos\left(d\left(z_i, m_j\right)\right)} \tag{4}$$

where $d(\cdot,\cdot)$ denotes a cosine similarity, which computes similarity measurement between the test query feature $z_i = f_{\theta_{enc}}(x_i)$ and the elements stored in the memory bank $\{m_j\}_{j=1}^{N}$. $\mathcal{A}(x_i)$ is normalized to $[0, 1]$. Ideally, the anomaly scores of anomalous images should be significantly larger than the scores from normal images. We also discard anomalous trained patterns in a memory bank as they can lead to adverse effects if anomalous prototypical patterns are similar to learned abnormal patterns.

3 Experimental Results

Datasets and Repartition. We validated our proposed method for classification of normal versus abnormal X-ray scans using two challenging public X-ray datasets, i.e., the NIH clinical center chest X-ray dataset [21] and the MURA (musculoskeletal radiograph) dataset [14]. In NIH dataset [21], each radiographic image is assigned with diagnostic labels corresponding to 14 cardiothoracic or pulmonary diseases. We combine all CXRs with at least one of these 14 diseases into an aggregate abnormal class. For a fair comparison with [20], we followed the same train, validation, and test subsets as in [20] so there was no patient ID overlap among the subsets. The MURA dataset [14] contains upper limb X-rays images labeled whether they contain anomaly or not. This dataset is composed of seven classes of body parts: finger, hand, wrist, forearm, elbow, humerus, and shoulder. There are a total of 40'005 X-ray images from 11'967 unique patients. We present a preprocessing pipeline, including the X-ray image carrier detection and unsupervised body part segmentation, using hysteresis thresholding by producing a binary mask (see Fig. 3). The splitting of the MURA dataset has been done based on the patient ID and the body part. This implies that all images of a specific body part from a given patient will be present in the same set. The patient's body parts are grouped into normal, abnormal, and mixed (meaning there are both normal and abnormal X-rays for that body part of a patient). The train set is composed mainly of normal samples (50% of all the patient's body part and 95% of normal samples) with few abnormal samples (5%). The remaining normal, abnormal, and mixed samples are equally split between the validation and the test sets (see Fig. 4).

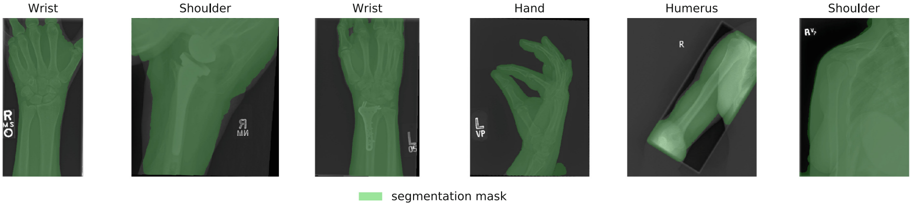


Fig. 3. Examples of segmentation results of the musculoskeletal X-rays.

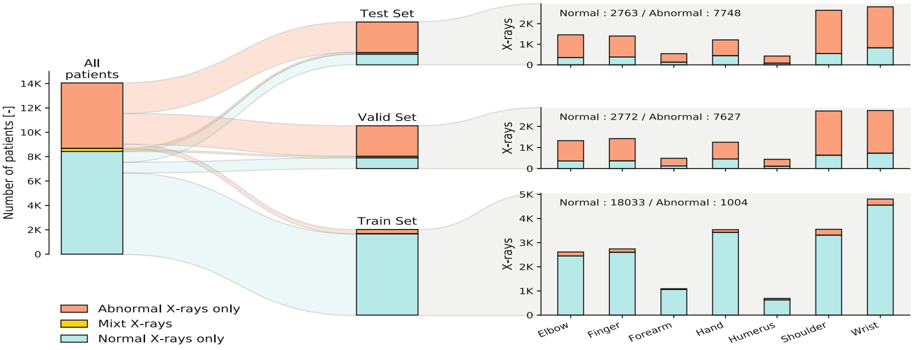


Fig. 4. A visual summary of the applied data split scheme on the MURA dataset.

Implementation Details and Evaluation Metrics. For the NIH dataset, we base our network architecture on the U-Net [15], consisting of a 6-layer convolutional encoder network and a 6-layer up-sampling convolutional decoder network without skip connections (both have batch normalization and leaky ReLU after each layer). The last encoder output features are projected to a 200-dimensional space, and L2 normalized. We use Adam optimizer, ($\beta_1 = 0.5, \beta_2 = 0.999$) and with a base learning rate of 0.0001. We pre-train the network for 50 epochs. Then, we train the network progressively with $\mathcal{L}_{\text{salad}}$ in 10 rounds with 50 epochs per round. The images were resized to 256×256 and we set $\tau = 0.1$, $\lambda = 0.25$, $t = 0.5$ and $K = 100$, respectively. The batch size was also set to 16. These optimum values are determined experimentally. We apply a slightly different experimental setup on the MURA dataset, and we replace the encoder and decoder with a ResNet-18 [11] and a mirrored ResNet-18, respectively. Besides, the musculoskeletal X-ray images are resized and padded so that their major axis is 512 pixels long while keeping the aspect ratio. We adopt area under the ROC curve (AUC) and Area Under Precision-Recall Curve (AUPRC) as our evaluation metrics.

Comparison with SOTA Unsupervised Methods. Figure 5 shows that our method significantly outperforms all recent anomaly detection methods, whether trained in an unsupervised mode, including OCGAN [20], Deep SVDD [16] and Deep Autoencoder (DAE) or with a self-supervised fashion (Geometric [7]). Although we use a few labeled anomalous data, the label information is not incorporated into our training, and we discard anomalous prototypical vectors for anomaly score calculation. Our method also achieves better performance compared to (GAN-GP) [9], where we replace GAN objective in [20] with gradient penalty (Fig. 6a).

Comparison to Methods that Use a Small Pool of Labeled Anomalies. To establish competing methods, we compare our method with the state-of-the-art semi-supervised method, Deep SAD [17], with the same data splitting and network bottleneck as outlined above. In addition, we train the supervised classifier, ResNet-18 [11] on the binary cross-entropy loss. We observed that our method surpasses other competing methods on two test sets (see Fig. 5). Semi-supervised and supervised approaches suffer from overfitting to previously seen anomalies at training while our self-supervised method generalizes well to unseen imbalanced anomalies (Fig. 6b).

Ablations. We conduct a series of ablation studies to justify the effectiveness of our contributions by comparing our full model with the following alternatives, using: 1. A Memory-based Deep Autoencoder (MemDAE) by turning off the proposed loss terms ($\mathcal{L}_{\text{agg}}$, $\mathcal{L}_{\text{ss}}$) and without using anomalous samples (Table 1 and Fig. 6a); 2. Our method without the loss term $\mathcal{L}_{\text{agg}}$; 3. Our method without the loss term $\mathcal{L}_{\text{mse}}$. We observed that our method trained with each of the

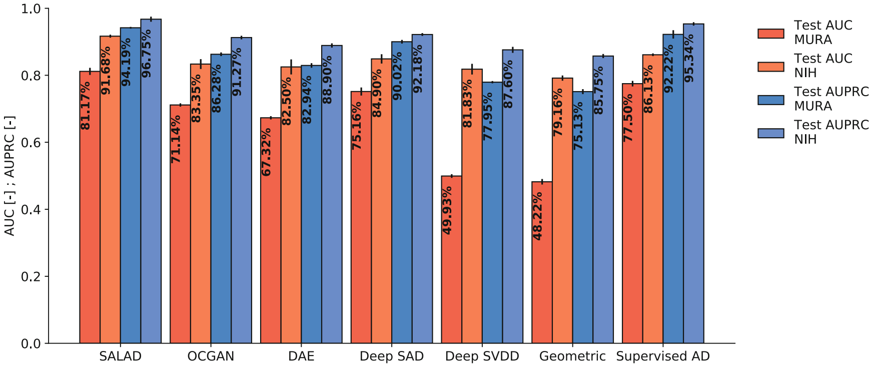


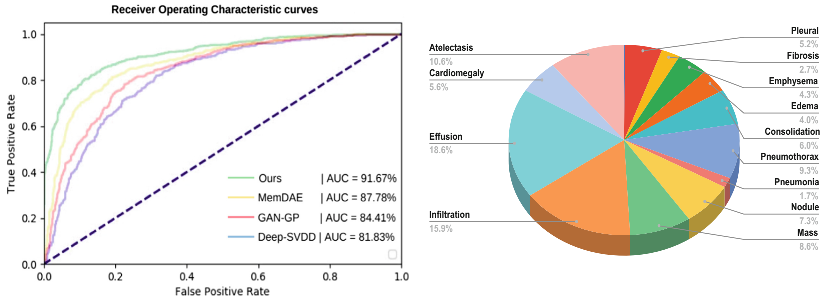
Fig. 5. AUC and **AUPRC.** Comparison of different anomaly detection methods. The bar height represents the mean (AUC or AUPRC) over four replicates of training, while the error bar is a 95% confidence interval computed as 1.96 std.

proposed loss terms, resulting in a notable performance gain over all the metrics, e.g., a gain of about 5.8% in AUC, compared to MemDAE on the MURA dataset (Table 1). Nevertheless, our baseline method (MemDAE) without anomalous samples, which is trained solely with MSE loss, outperforms all previous anomaly detection methods. We also conduct sensitivity analysis to investigate the effect of included labeled anomalies during training on final performance. To do so, we increased the ratio of known anomalous samples up to 15% and observed that our method is not very sensitive to an anomalous ratio (Table 1). This can be explained by the fact that SALAD does not require label information. Instead, it uses anomalous samples to have a better separation of prototypical patterns.

Table 1. Evaluation of the proposed approach and baselines. Ablation study for anomalous ratio (AR) and the loss terms used in training.

Method	AR (%)	NIH		MURA	
		AUC	AUPRC	AUC	AUPRC
MemDAE	0	0.8778 ± 0.0072	0.9198 ± 0.0024	0.7660 ± 0.0102	0.8823 ± 0.0035
SALAD w/o $\mathcal{L}_{\text{mse}}$	5	0.8693 ± 0.0095	0.9121 ± 0.0032	0.7765 ± 0.0094	0.8892 ± 0.0085
SALAD w/o $\mathcal{L}_{\text{agg}}$	5	0.8824 ± 0.0105	0.9332 ± 0.0062	0.7923 ± 0.0061	0.9128 ± 0.0026
SALAD (ours)	5	0.9091 ± 0.0069	0.9552 ± 0.0049	0.8117 ± 0.0028†	0.9419 ± 0.0075†
SALAD w/o $\mathcal{L}_{\text{mse}}$	10	0.8739 ± 0.0071	0.9287 ± 0.0024	0.7882 ± 0.0104	0.8926 ± 0.0031
SALAD w/o $\mathcal{L}_{\text{agg}}$	10	0.8913 ± 0.0056	0.9426 ± 0.0035	0.8012 ± 0.0042	0.9196 ± 0.0023
SALAD (ours)	10	0.9167 ± 0.0031†	0.9674 ± 0.0051†	0.8195 ± 0.0087	0.9502 ± 0.0085
SALAD w/o $\mathcal{L}_{\text{mse}}$	15	0.8774 ± 0.0095	0.9290 ± 0.0023	0.7904 ± 0.0025	0.8985 ± 0.0082
SALAD w/o $\mathcal{L}_{\text{agg}}$	15	0.8975 ± 0.0105	0.9494 ± 0.0055	0.8084 ± 0.0014	0.9214 ± 0.0103
SALAD (ours)	15	0.9189 ± 0.0076	0.9705 ± 0.0011	0.8245 ± 0.0017	0.9582 ± 0.0108

† Final models used for comparison against SOTA methods.



(a) ROC curves

(b) The statistics of the NIH set

Fig. 6. (a) ROC curves comparison performances on the NIH dataset. (b) The NIH dataset statistics used in our experiments.

4 Conclusion and Future Work

In this work, we proposed SALAD, a self-supervised aggregation based learning framework for X-ray anomaly detection. This paper's novelty lies in jointly deep representation learning of X-ray images as well as aggregation criterion to distill out anomalous data. We use progressive training to enforce consistency between similar data samples in the embedding space to facilitate the formation of prototypical normal patterns. Hence, abnormal X-ray samples appear less likely to be represented by the normal learned patterns. SALAD achieves state-of-the-art anomaly detection results across all tested learning regimes, including unsupervised methods and those trained with small amounts of labeled data. As future work, we envision the broad application of our approach across different image modalities and beyond anomaly detection where the annotation is very costly, e.g., unsupervised domain adaptation.

References

1. Alaverdyan, Z., Jung, J., Bouet, R., Lartizien, C.: Regularized siamese neural network for unsupervised outlier detection on brain multiparametric magnetic resonance imaging: application to epilepsy lesion screening. Med. Image Anal. **60**, 101618 (2020)
2. Baur, C., Wiestler, B., Albarqouni, S., Navab, N.: Fusing unsupervised and supervised deep learning for white matter lesion segmentation. In: International Conference on Medical Imaging with Deep Learning, pp. 63–72 (2019)
3. Breunig, M.M., Kriegel, H.P., Ng, R.T., Sander, J.: Lof: identifying density-based local outliers. In: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, pp. 93–104 (2000)
4. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations. arXiv preprint arXiv:2002.05709 (2020)
5. Davletshina, D., et al.: Unsupervised anomaly detection for x-ray images. arXiv preprint arXiv:2001.10883 (2020)

6. Gidaris, S., Singh, P., Komodakis, N.: Unsupervised representation learning by predicting image rotations. arXiv preprint arXiv:1803.07728 (2018)
7. Golan, I., El-Yaniv, R.: Deep anomaly detection using geometric transformations. In: Advances in Neural Information Processing Systems, pp. 9758–9769 (2018)
8. Gong, D., Liu, L., Le, V., Saha, B., Mansour, M.R., Venkatesh, S., Hengel, A.V.D.: Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1705–1714 (2019)
9. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C.: Improved training of wasserstein gans. In: Advances in Neural Information Processing Systems, pp. 5767–5777 (2017)
10. 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)
11. 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)
12. Liu, F.T., Ting, K.M., Zhou, Z.H.: Isolation forest. In: 2008 Eighth IEEE International Conference on Data Mining, pp. 413–422. IEEE (2008)
13. Norlander, E., Sopasakis, A.: Latent space conditioning for improved classification and anomaly detection. arXiv preprint arXiv:1911.10599 (2019)
14. Rajpurkar, P., et al.: Mura: Large dataset for abnormality detection in musculoskeletal radiographs. arXiv preprint arXiv:1712.06957 (2017)
15. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds.) MICCAI 2015. Lecture Notes in Computer Science, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
16. Ruff, L., et al.: Deep one-class classification. In: International Conference on Machine Learning, pp. 4393–4402 (2018)
17. Ruff, L., et al.: Deep semi-supervised anomaly detection. In: International Conference on Learning Representations (2020). https://openreview.net/forum?id=HkgH0TEYwH
18. Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G.: Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: Niethammer, M., et al. (eds.) IPMI 2017. LNCS, vol. 10265, pp. 146–157. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59050-9_12
19. Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R.C.: Estimating the support of a high-dimensional distribution. Neural Comput. **13**(7), 1443–1471 (2001)
20. Tang, Y.X., Tang, Y.B., Han, M., Xiao, J., Summers, R.M.: Abnormal chest x-ray identification with generative adversarial one-class classifier. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 1358–1361. IEEE (2019)
21. 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)
22. Wu, Z., Xiong, Y., Yu, S.X., Lin, D.: Unsupervised feature learning via non-parametric instance discrimination. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3733–3742 (2018)

23. Zhai, S., Cheng, Y., Lu, W., Zhang, Z.: Deep structured energy based models for anomaly detection. In: International Conference on Machine Learning, pp. 1100–1109 (2016)
24. Zhuang, C., Zhai, A.L., Yamins, D.: Local aggregation for unsupervised learning of visual embeddings. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 6002–6012 (2019)
25. Zong, B., et al.: Deepautoencoding gaussian mixture model for unsupervised anomaly detection (2018)