AJBAR Vol. 3(4), August 2024: 136-145, ISSN: 2811-2881

Arid Zone Journal of Basic and Applied Research

Faculty of Science, Borno State University
Maiduguri, Nigeria

Journal homepage: https://www.azjournalbar.com

Research Article

A Data-Efficient Deep Learning Approach for Lung Nodule Segmentation

*Murtala Dattijo Makama[1], Abdulkarim Salihu Aish[1], Ibrahim Dattijo Makama[2]
[1]National Information Technology Development Agency (NITDA)
[2]National Open University of Nigeria (NOUN)
***Corresponding author's Email**: mdattijo@nitda.gov.ng, doi.org/10.55639/607.02010006

ARTICLE INFO:

Keywords:
Self-supervised learning,
Deep learning,
Medical imaging,
Segmentation,
Lung CT scans

ABSTRACT

Medical imaging is undeniably a very important field in the provision of effective health care. It allows for early detection, diagnosis, and treatment of different kinds of diseases. On the other hand, preparation of labeled data from the biomedical domain poses difficulties, since it requires heavy privacy considerations and the expert radiologist to perform manual labeling of that data. Due to that reason, efforts to utilize advancements in deep learning algorithms are often hampered. Generating expert annotation of medical images at scale is difficult, costly, and time-consuming. This challenge is even greater in developing countries, especially in Africa, where there is a massive shortage of radiologists to annotate these images. Another related issue in medical imaging is the scarcity of datasets, particularly in Africa, where there is a severe shortage of digital imaging modalities. This becomes more obvious when studying a particular disease, for instance. Hence, the scarcity of large, well-labelled datasets is the main constraint for deep learning applications in medical imaging. To solve this problem, we devise a self-supervised learning approach. The proposed method used inpainting as the pretext task and U-Net for segmentation. The pre-trained network was then fine-tuned on a varied sample of annotated datasets. Our results performance improvement with 50% of labeled data compared to supervise learning. This approach has the potential to significantly improve the applicability of deep learning for medical imaging in resource-constrained settings.

Corresponding author: Murtala Dattijo Makama, **Email:** mdattijo@nitda.gov.ng
National Information Technology Development Agency (NITDA)

INTRODUCTION

Lung cancer is one of the most common and deadly cancers in the world (Surveillance Research Program, National Cancer Institute, 2024). The World Health Organization reported that lung cancer in low- and middle-income countries is the leading and the second-leading cause of cancer deaths among males and females, respectively (Bray *et al.*, 2018). According to the American Cancer Society, the 5-year survival rate of lung cancer is 17.8%, much lower compared to that of breast cancer at 90.35% (Surveillance Research Program, National Cancer Institute, 2024). The pulmonary nodules represent a potential symptom of lung cancer, for which detection at an early stage is crucial for the improvement of survival rates among patients (Yablonskii *et al.,* 2020). Early detection means an enhanced chance of survival if the nodules are detected at an early stage (Yablonskii et al., 2020).

Medical images can play an important role in health care for patients by preventing diseases, which can be done through their early detection and diagnosis, even treatment (Malhotra *et al.,* 2022). CT is the major imaging technique applied in diagnosing and detecting lung cancers. However, the detection of lung nodules is a very challenging and time-consuming task, even for expert radiologists. To that end, computer vision techniques, with deep learning being particularly prominent, have been increasingly used for lung nodule detection with considerable success recently (Greenspan et al., 2016). However, most of such approaches require a large number of well-structured and hand-labelled data, which is indeed a very critical challenge in many applications, especially in the medical domain. Hence, overcoming these issues, the work presented now is related to the elaboration and evaluation of a self-supervised approach for lung nodule segmentation. The proposed technique in this respect relies on the use of inpainting as a pretraining task (Pathak *et al.*, 2016), followed by fine-tuning of the U-Net model (Ronneberger *et al.,* 2015) using limited available annotated data. This thereby learns to make effective use of unlabelled data for better performance in segmentation tasks with a minimum burden of extensive manual annotations.

SSL provides a strong approach to utilizing the enormous, easily available amount of data in an unlabelled manner (Ozbulak *et al.,* 2023). Since SSL automatically generates supervisory signals from the unlabelled data itself, it enables the learning of meaningful representations from the intrinsic structures and relationships within the data. It hence promises significant advances in medical imaging, especially in resource-constrained settings where annotated datasets are at a premium (Krishnan *et al.,* 2022).

The main motivation remains the same: high-quality labelled datasets can be very expensive and time-consuming to create, whereas unlabelled data is continuously generated (Weng, 2019). SSL addresses this challenge by generating supervisory signals from the unlabelled data and provides a way for the model to learn meaningful representation from the intrinsic structures and relationships of the data (Chen *et al.,* 2019). This could accelerate medical artificial intelligence development and increase the applicability of deep learning to medical imaging, especially in resource-constrained environments.

RELATED WORKS

Self-supervised learning has recently emerged as a solution to utilize unlabelled datasets during neural network training (Spathis *et al.,* 2022). The advantage of self-supervised methods is the ability to learn useful representations without a need for manually labelled examples. It enables incorporating unlabelled examples to perform network pre-training (Ozbulak *et al.,* 2023). While deep learning, and specifically convolutional neural networks, has had great success recently, this may be partly due to exponential increases in available annotated data. For highly specialized domains, such as medical imaging and radiology, it is much harder to come across annotations (Greenspan *et al.,* 2016). Thus, different strategies have been proposed that differ in the kinds of used pretext tasks and objective functions. There are generative methods and contrastive methods.

The generation-based methods involve learning representations by networks that learn to generate some content. The authors of (Dosovitskiy *et al.,* 2015) propose an architecture that restores the content of the randomly chosen patch removed from the image. In (Zhang *et al.,* 2016), the authors propose image colourization as a form of the pretext task. Images in the dataset are

converted to greyscale, and then the task of the network is to restore full information about the colour. Context-based pretext tasks rely on the semantic and spatial information within the image. Furthermore, the authors of (Selvan *et al.*, 2020) used a generative approach to self-supervised learning to segment high-opacity chest X-rays (CXRs) from normal CXRs. In contrast, approaches based on contrastive losses have recently shown great potential (Chen *et al.*, 2019). This family of methods involves applying modification to the image. The network is trained in such a way as to make representations of two different views of the image similar while making those views dissimilar to other examples in the dataset. This approach makes the representations invariant under transformation applied to the images (Doersch *et al.*, 2015). It was shown that the quality of the learned representations benefits from such a form of pre-training.

The authors of (Chen *et al.*, 2020) propose a framework called "SimCLR" for modelling the above problem in a self-supervised manner. It blends the concept of Contrastive Learning with a few novel ideas to learn visual representations without human supervision (Chen *et al.*, 2019). The framework takes an image and applies random transformations to it to get a pair of augmented images xi and xj. Each image in that pair is passed through an encoder to get representations. Then a non-linear, fully connected layer is applied to get representations zi and zj, respectively. The task is to maximize the similarity between these two representations zi and zj for the same image. The authors of (Chaitanya *et al.*, 2020) adopted and extended the works of (Chen *et al.*, 2020) in the medical domain by introducing a local version of the contrastive loss to learn distinctive representations of local regions that are useful for per-pixel segmentation.

In terms of feature learning, another self-supervised research proposed, which is named context encoders. Feature learning by inpainting (Pathak *et al.*, 2016) is trained to fill in a missing piece in the image. Reconstructing the original input allows learning meaningful latent representation. But the method comes with a shortcoming—it changes the image intensity distribution. Thus, the resulting images belong to another domain, and the learned features may not be useful for images in the original domain.

The authors of (Chen *et al.*, 2019) propose a novel strategy for self-supervised which they term context restoration. The paper validates the context restoration strategy in three common problems in medical imaging: classification, localization, and segmentation. The method is straightforward that select random two isolated small patches in a given image and swap their context. Repeat these operations T times, in which the intensity distribution is still preserved, but its spatial information is altered.

In (Gidaris *et al.*, 2018) the proposed network is called RotNet. The dataset used during pretext task training is created by performing random rotations to images. The network task is to predict the rotation angle. Although it can be easily implemented, the method provides a significant improvement. DeepCluster, proposed in (Caron *et al.*, 2018) uses a clustering algorithm to group representations into clusters. Then, those groups are utilized during supervised training. Thus, the training involves alternating between clustering the representations into a group and training the network to predict which group the image belongs to.

METHODOLOGY

Our proposed self-supervised method for pulmonary nodule segmentation involves two key stages: pretext task training and downstream task training. To improve the clarity and organization of this section, the methodology is structured into three distinct subsections: Pretext Task, Downstream Task and Data Augmentation.

Pretext Task

The pretext task training is performed to learn visual features that can be utilized in the downstream task (Ozbulak *et al.*, 2023). This stage uses an unlabelled dataset, where the common pipeline involves applying transformations to the images and training the network to predict these transformations (Weng, 2019). For our study, we employed inpainting as the pretext task (Pathak *et al.*, 2016), where parts of the lung CT scan patches (128x128x1) were corrupted and then fed into a modified U-Net autoencoder (Ronneberger *et al.*, 2015). The autoencoder reconstructs the missing pixels, and the reconstruction loss (L2) is computed by comparing the generated image with the original. This task helps the network learn meaningful latent features, crucial for effective segmentation. This is

considered stage 1 of the self-supervised process, as shown in Fig. 1.

Downstream Task

After completing the pretext task, the learned weights are used to initialize the parameters for the downstream task, which is the primary task of pulmonary nodule segmentation. Similar to the pretext task, the U-Net architecture is utilized here. The input images and their corresponding masks are processed through the network to produce segmented outputs. The binary cross-entropy loss function is used during training, with the initialized weights from the pretext task providing a strong starting point for the segmentation task.

Data Augmentation

Data augmentation has been used here to enhance the generalization performance of the deep neural network by incorporating more diversity into the training data (Xu *et al.*, 2020). Several augmentation techniques used in this study, such as rotation, horizontal flipping, vertical flipping, and combined flipping, were utilized.

Each of these augmentation techniques contributes its share to increasing model performance by making the data more variant.

- Rotation would enable the model to recognize the nodules from a different angle, which is important because nodules may appear in many orientations within the lungs.
- Horizontal and vertical flipping will avoid orientation bias of the model for nodules, enabling it to generalize well for any patient's scan.

We have excluded stretching and skewing since they may distort the shape by not uniformly transforming the nodules. This is very important for correct segmentation, as its shape could easily be changed to mislead the model during training and decrease the accuracy of the segmentation job. We select the augmentation techniques in such a way that they maintain the structural integrity of the nodules, hence making sure that the model will learn to segment the lung nodules in an orientation and positioning variance robust manner.

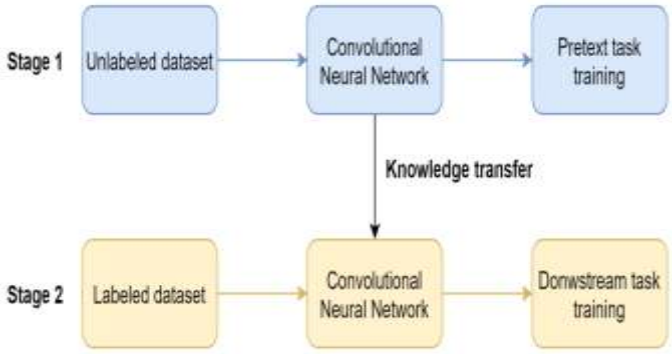
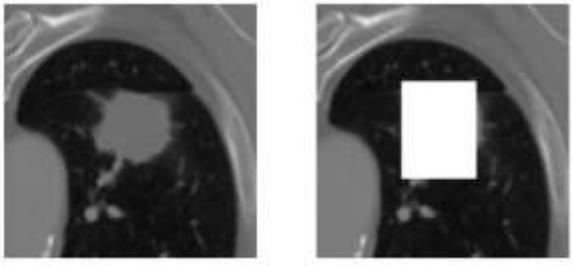


Figure 1: The general pipeline of self-supervised learning. An unlabelled dataset is used to pre-train the network (Stage 1), that is then used to train the network on a labelled dataset (Stage 2)



(a) No augmentation (b) Inpainting

Figure 2: 2D slices of CT scans of lung legions with (a) and without data augmentation (b).

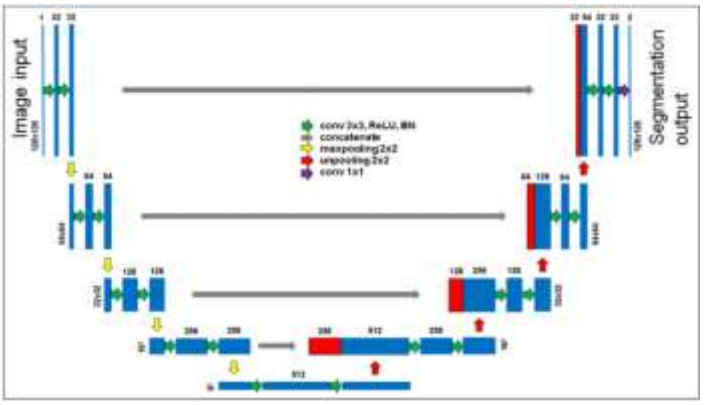


Figure 3: Schematic diagram of 2D U-Net: a CT image of 128 x 128 pixels was input, down-sampled to 8 x 8, and then up-sampled.

EXPERIMENT AND RESULTS

Data

The dataset used in this research consists of 520 cases extracted from the publicly available data from the Lung Image Database Consortium image collection (LIDC/IDRI) (Armato *et al.*, 2011). It is the largest publicly available reference database of chest CT scans (Armato *et al.*, 2011). It consists of diagnostic and lung cancer screening computed tomography (CT) scans with marked-up annotated lesions. CT images from each scan are associated with an XML file recording nodule reports and diagnosis reports of the image annotation process from four experienced thoracic radiologists. Each LIDC/IDRI annotation was created by a two-phase reading process. In the first blinded reading phase, suspicious lesions were independently annotated by four experienced thoracic radiologists as non-nodule >3mm, nodule >3mm, or non-nodule (any other pulmonary abnormality). In the second non-blind reading phase, the blinded results of all other radiologists were revealed to each radiologist, who then decided to accept or reject each annotation. No consensus was forced.

Pre-processing

In the proposed method, all the input images were re-scaled to 128x128. Before dividing it, we apply augmentation as described in the methodology. After augmenting the data, we have 520 positive and negative patches, which are divided into training and testing sets. The training set consists of varied cases ranging from 50 to 500 and the text set consists of a fixed 20 cases.

Evaluation

The dice coefficient loss is selected as the loss function (Han et al., 2020). The dice coefficient, as in (1) is often used in medical image segmentation (Shi *et al.*, 2022).

$$\text{Dice Coefficient} = 2 * |X \cap Y| / (|X| + |Y|) \quad (1)$$

It is often used to calculate the similarities between two cases.

Experimental setup

A series of experiments were done to evaluate the performance of the various strategies used. We first tried supervised learning with a varied number of training sets ranging from 50 to 500 cases. We then perform the self-supervised learning approach using inpainting as proposed by (Pathak *et al.*, 2016). We show the Dice score for each of the models in Table 1, and the segmentation results produced by the model are shown in Fig. 5. The proposed model is based on the state-of-the-art segmentation model, U-Net (Ronneberger *et al.*, 2015). The deep learning framework used in building the model is TensorFlow (Abadi *et al.*, 2015). Before performing the proposed approach of self-supervised learning, we performed supervised learning with a reduced number of training cases, and the results are shown in Table 1. We split each dataset into a training set and a test set, each consisting of 2D scan slices with their corresponding segmentation labels. The model was pre-trained as described in the methodology. The input layer contained 128x128 pixels with 1 channel. We performed 2D convolution (Conv2D) by applying a 2x2 filter to the input, in which the 'same' padding filled the perimeter of the input with zeros compensating for the reduction in size of the image produced by the filter. Max pooling employing

MaxPooling2D, by which the maximum value is selected from each region and then compressed, was performed 2-dimensionally to reduce pixel size. Conv2DTranspose is an operation that is the opposite of that used for Conv2D, whereby pixel size is increased using a 2x2 pixels filter. Batch Normalization (BN) is a process by which biased output distribution, obtained from the previous layer, is corrected. Dropout is a process by which a portion of the units are randomly deactivated. Concatenation is a process applied to connect the input arrays. Rectified linear unit (ReLU) was used for the input and hidden layers, and sigmoid for the output layer as the activation function. The pre-trained model is then fine-tuned on varied percentages of labelled data ranging from 50 to 500 cases from the training set for the downstream task of pulmonary nodule segmentation using cross-entropy loss. Stochastic Gradient Descent with momentum algorithm is used as the optimizer, with learning rate set to 0.001, momentum coefficient set to 0.9, and batch size set to 32. The test set consists of 20 fixed cases. It is used for training the algorithm, whereas the test set is employed to determine the performance of the algorithm.

Table 1: Models' performance on the dataset

Models	Training Size (cases)	Testing Size (cases)	Dice Score
Vanilla U-Net	50	20	0.5050
	100	20	0.6044
	150	20	0.7622
	200	20	0.8355
U-Net with pretext training	50	20	0.6063
	100	20	0.6710
	150	20	0.7830
	200	20	0.8518

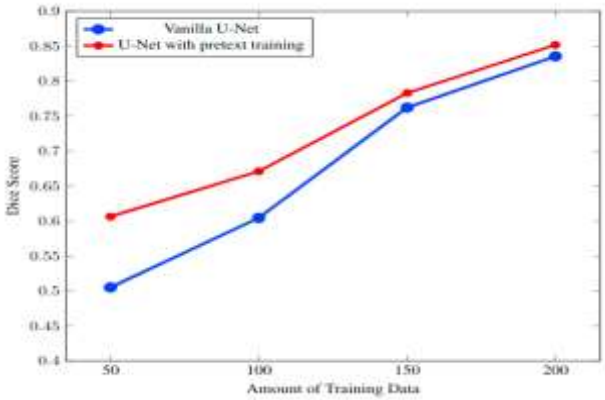
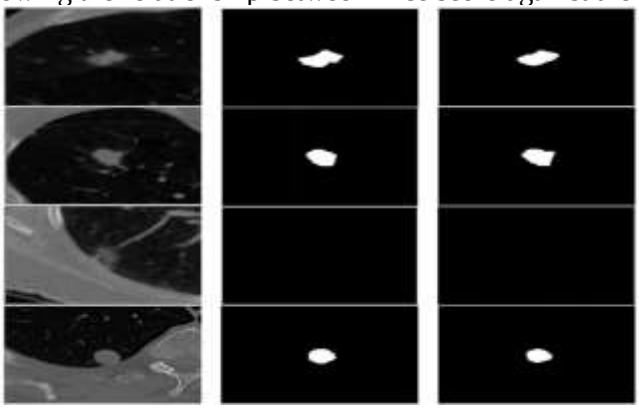


Figure 4: Line plot showing the relationship between Dice score against the amount of training data



(a) (b) (c)

Figure 5: Segmentation results (a). Original 2D CT image (b) Ground truth. (c). Result.

DISCUSSION

In this work, we demonstrate that the use of the self-supervised learning method of inpainting during the pretext task, supplemented with extensive data augmentation, significantly enhances the performance of the model on lung nodule segmentation. On the other hand, the proposed method obtains the same accuracy of nodule segmentation as fully supervised learning models even when trained with 50% labeled data. This is particularly important in medical imaging, where it is often difficult and expensive to obtain large, well-annotated datasets (Krishnan *et al.*, 2022).

Analysis of Results

The Dice coefficient plotted, as the performance of the model for segmenting, indicates that self-supervised pretraining effectively captured vital features in the lung nodules. Such are then fine-tuned in the subsequent segmentation task. Inpainting helps the model learn robust features by reconstructing the missing parts of an image. Besides being very important, this intrinsic capability enables the nodules to be detected. Apart from that, rotation and flipping features in data augmentation will enhance the generalizing ability of the model across various variations that may appear in the dataset. This ensures that the segmentation isn't biased towards any particular orientation or position of the nodules.

These results suggest that this approach reduces the dependency not only on large annotated datasets but also increases model robustness, enhancing its adaptiveness to changes in the input data. The capability of high performance with less labeled data is highly desirable in resource-constrained settings where access to annotated medical images is limited (Krishnan *et al.*, 2022).

Implications

The implications of these findings are significant for clinical settings. By reducing the need for extensive manual annotation, the proposed method can greatly improve radiologists' efficiency, allowing them to focus on more complex cases while the model handles routine segmentation tasks. This automation could also expedite the diagnostic process, enabling quicker detection and treatment of lung cancer, which is critical for improving patient outcomes.

Moreover, the method's adaptability suggests that it could be integrated into large-scale lung cancer screening programs, providing a reliable tool for early detection of pulmonary nodules. The ability to maintain high accuracy with reduced labeled data also means that the model can be effectively used in regions with limited medical resources, contributing to global health efforts in combating lung cancer.

Limitations

Despite the success of such promising results, there are several limitations that need consideration about the proposed method. First, most of the superior performance depends on some particular types of data augmentation, which may not be generalized for all medical image tasks (Xu *et al.*, 2020). For example, rotation and flipping are informative in the segmentation of lung nodules, but different augmentation is desired for other kinds of medical images in order to preserve the critical features.

Another limitation is the performance of the model, which may be different when various data sets or imaging modalities are used. The model was developed and checked on one data set of lung CTs, and its performance may differ when dealing with other types of medical images.

Another limitation of this study is that, in the self-supervised learning phase, only one pretext task-inpainting-was tried out. While that worked rather well, it is probably not the best pretext task available (Ozbulak *et al.*, 2023). Alternative pretext tasks, including but not limited to contrastive learning and generative modeling, may yield further improvements in the model.

Future Research

This work points to several interesting avenues for further studies. One such avenue of interest could be an investigation of applying the proposed method to brain MRI segmentation, detection of breast cancer in mammograms, or analysis of retinal images. All these tasks have their specific complications; adaptation of the self-supervised approach for image segmentation in these tasks may bring significant benefits for diagnostics in many medical spheres.

Another interesting avenue of future research lies in how multiple pretext tasks might be combined during the self-supervised learning phase to yield even better feature extraction. For instance, such work might combine the inpainting task (Pathak et al., 2016) with other tasks, such as jigsaw puzzles or colorization, toward an even more complete understanding of image structure that results in even better segmentation.

Further, increasing the dataset in the number and variability of lung CT scans from different populations and scanning conditions would help investigate generalizability. Performance testing on data across a range of demographic groups, scanner types, and levels of image quality would provide highly useful information about its robustness and reliability in the real clinical environment.

The model can be further tuned with respect to its integration with other advanced techniques, such as transfer learning and federated learning, which would decrease its need for labeled data even more, besides extending its applicability into decentralized healthcare environments.

CONCLUSION

Obtaining a high-quality dataset in the medical domain remains a huge challenge, especially in low-resource settings such as Africa (Krishnan et al., 2022). Self-supervised learning can be beneficial in this regard. Our end-to-end neural network automatically learned nodule-sensitive features from the 2D CT images to improve the segmentation performance. This was made possible by the self-supervised learning technique of inpainting (Pathak et al., 2016) which helps the network learn useful latent features as it reconstructs the missing pixels in the image during the pretext task. The proposed method can support the research on automated lung nodule segmentation as well as offer a powerful and effective tool for computer-aided diagnosis of lung cancer, where accurate segmentation of pulmonary nodules is essential. Given the accuracy of our algorithm, our automated nodule detection tool has the potential to be a powerful clinical tool that can be implemented in clinical practice for the accurate detection of nodules in lung cancer screening projects. This tool will not only enhance the manual identification of pulmonary nodules but will also reduce the reading time when used as an assisting tool to augment the radiologist's workflow. Overall, we believe that the proposed initialization, combined with a data augmentation technique such as inpainting (Pathak *et al.*, 2016) provides a simple toolbox for vastly improving performance in segmentation tasks in medical imaging, especially in clinically relevant low-resource settings.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... & Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Armato, S. G., McLennan, G., Bidaut, L.,McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., ... & Croft, B. Y. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. Medical Physics, 38(2), 915–931.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians, 68(6), 394–424.

Caron, M., Bojanowski, P., Joulin, A., & Douze, M. (2018). Deep clustering for unsupervised learning of visual features. European Conference on Computer Vision.

Chaitanya, K., Erdil, E., Karani, N., & Konukoglu, E. (2020). Contrastive learning of global and local features for medical image segmentation with limited annotations. arXiv preprint arXiv:2006.10511.

Chen, L., Bentley, P., Mori, K., Misawa, M., Fujiwara, M., & Rueckert, D. (2019). Self-supervised learning for medical image analysis using image context restoration. Medical Image Analysis, 58, 101539.

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. arXiv preprint arXiv:2002.05709.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A

large-scale hierarchical image database. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Doersch, C., Gupta, A., & Efros, A. A. (2015). Unsupervised visual representation learning by context prediction. arXiv preprint arXiv:1505.05192.

Dosovitskiy, A., Fischer, P., Springenberg, J. T., Riedmiller, M., & Brox, T. (2015). Discriminative unsupervised feature learning with exemplar convolutional neural networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(9), 1734–1747.

Gidaris, S., Singh, P., & Komodakis, N. (2018). Unsupervised representation learning by predicting image rotations. International Conference on Learning Representations.

Greenspan, H., Van Ginneken, B., & Summers, R. M. (2016). Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE Transactions on Medical Imaging, 35(5), 1153–1159.

Han, T., Nunes, V. X., De Freitas Souza, L. F., Marques, I. C. L., Silva, M. A. A. F., Junior, P. P. R., & Filho, J. (2020). Internet of medical things—based on deep learning techniques for segmentation of lung and stroke regions in CT scans. IEEE Access, 8, 71117–71135.

Hendrycks, D., & Dietterich, T. (2019). Benchmarking neural network robustness to common corruptions and perturbations. Proceedings of the International Conference on Learning Representations.

Kayal, S., Chen, S., & de Bruijne, M. (2020). Region-of-interest guided supervoxel inpainting for self-supervision. arXiv preprint arXiv:2005.12052.

Krishnan, R., Rajpurkar, P., & Topol, E. J. (2022). Self-supervised learning in medicine and healthcare. Nature Biomedical Engineering, 6(12), 1346–1352.

Li, F.-F., Fergus, R., & Perona, P. (2004). Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Malhotra, P., Gupta, S., Koundal, D., Zaguia, A., & Enbeyle, W. (2022). Deep neural networks for medical image segmentation. Journal of Healthcare Engineering, 2022.

Ozbulak, U., Lee, H. J., Boga, B., Anzaku, E. T., Park, H., Messem, W. D., & Neve, J. (2023). Know your self-supervised learning: A survey on image-based generative and discriminative training. arXiv preprint arXiv:2305.14005.

Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., & Efros, A. A. (2016). Context encoders: Feature learning by inpainting. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U- Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 234–241.

Selvan, R., Dam, E. B., Detlefsen, N. S., Rischel, ., Sheng, K., Nielsen, M., & Pai, A. (2020). Lung segmentation from chest x-rays using variational data imputation. Proceedings of the International Conference on Machine Learning Workshop.

Shi, P., Duan, M., Yang, L., Feng, W., Ding, L., & Jiang, L. (2022). An improved U-Net image segmentation method and its application for metallic grain size statistics. Materials, 15(13), 4545.

Spathis, D., Perez-Pozuelo, I., Marques-Fernandez, L., & Mascolo, C. (2022). Breaking away from labels: The promise of self-supervised machine learning in intelligent health. Patterns, 3(2), 100410.

Surveillance Research Program, National Cancer Institute. (2024). SEER Explorer: An interactive website for SEER cancer statistics. Accessed at https://seer.cancer.gov/explorer/on.

Taleb, A., Loetzsch, W., Danz, N., Severin, J., Gaertner, T., Bergner, B., & Lippert, C. (2020). 3D self-supervised methods for medical imaging. arXiv preprint arXiv:2010.11465.

Tran, G. S., Nghiem, T. P., Nguyen, C. M., Luong, C. M., & Burie, J.-C. (2019). Improving accuracy of lung nodule classification using deep learning with focal loss. Journal of Healthcare Engineering, 2019.

Welinder, P., Branson, S., Mita, T., Wah, C., Schroff, F., Belongie, S., & Perona, P. (2010). Caltech-UCSD Birds 200. Technical Report CNS-TR-2010-001, California Institute of Technology.

Weng, L. (2019). Self-supervised representation learning. Lilianweng.github.io/lil-log. Retrieved from https://lilianweng.github.io/lil-log/2019/11/10/self-supervised-learning.html.

Xu, J., Li, M., & Zhu, Z. (2020). Automatic data augmentation for 3D medical image segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.