Domain Generalization for Pulmonary Nodule Detection via Distributionally-Regularized Mamba

Tianzhong Lan[1], Nan Chen[2], Zhang Yi[1], Xiuyuan Xu[1](✉), and Min Zhu[1](✉)

[1] College of Computer Science, Sichuan University, Chengdu 610065, China
[2] West China Hospital, Sichuan University, Chengdu 610044, China
{zhumin, xuxiuyuan}@scu.edu.cn

Abstract. Extending deep learning models to out-of-distribution (o.o.d.) data remains a persistent challenge, especially in domains like medical imaging with restricted data availability and limited data sharing. This challenge is particularly evident in pulmonary nodule detection, as the model struggles to distinguish nodules from the surrounding normal tissues across different data distributions. To address this issue, we propose a Distributionally Regularized Mamba Network (DRMNet). Inspired by Mamba, we propose a Feature-Augmented State-Space module that unifies pulmonary nodule features to effectively distinguish nodules from surrounding confounding tissues. Furthermore, a Region-Aware Distribution Alignment module is elaborately introduced to reduce disparities in feature distributions between domains. We construct a pulmonary nodule detection dataset, named Generalization for Pulmonary Nodule Detection (GPND), comprising diverse domains, including private and well-known public datasets. Extensive experiments conducted on GPND demonstrate that DRMNet outperforms state-of-the-art domain generalization methods. The code is available at `https://github.com/TzhongBoyyy97/DRMNet`.

Keywords: Pulmonary Nodule Detection · Nodule Dataset · Domain Generalization

1 Introduction

Lung cancer has a high mortality rate, but early CT screening can reduce deaths. Existing pulmonary nodule detection algorithms [8, 14, 17, 22, 23] overlook common out-of-distribution (o.o.d.) scenarios in clinical practice (see Fig. 1). To tackle domain shifts in medical imaging, domain generalization (DG) enhances model generalizability to *unseen* target domains by leveraging information from one or more source domains [19, 28]. DG approaches typically involve data augmentation [6, 24, 25], meta-learning [3, 5, 9], or learning domain-invariant features [2, 26, 27].

While existing methods show promise, limited research has been dedicated to DG in pulmonary nodule detection. Unlike other lesions, pulmonary nodules are

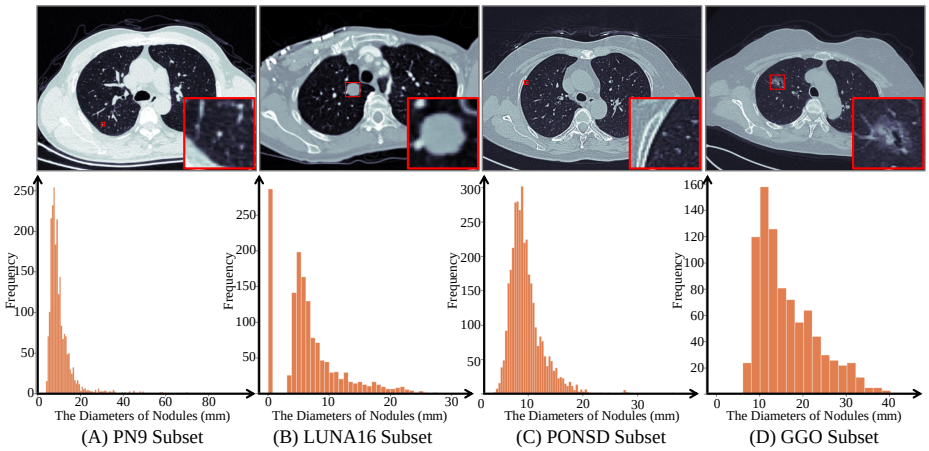

Fig. 1: Comparsion of PN9, LUNA16, PONSD and GGO. The first row represents some nodule samples, and the second row indicates the diameter distributions in four datasets. Four subsets introduce clear domain shifts.

intertwined with blood vessels and other normal tissues in imaging, making them difficult to identify. Therefore, capturing long-range contextual representations is essential for distinguishing spherical nodules from tubular vascular structures. Recently, Mamba [4] has demonstrated effectiveness in capturing long-range dependencies in tasks such as image classification and segmentation while avoiding excessive GPU memory usage like Transformers. This capability could enhance detection models by improving the extraction of domain-invariant features from pulmonary nodules.

In this paper, we introduce a Generalization for Pulmonary Nodule Detection (GPND) dataset, comprising publicly available subsets (LUNA16 [16] and PN9 [14]) and internal subsets (PONSD and GGO), each characterized by unique imaging qualities and nodule characteristics. We further propose a Distributionally Regularized Mamba Network (DRMNet), integrating a U-shaped architecture with Feature-Augmented State-Space (FASS) modules to capture both local fine-grained features of pulmonary nodules and long-range correlation between nodules and vasculars. Additionally, DRMNet incorporates a Region-Aware Distribution Alignment (RADA) module to constrain the latent feature distribution to a predefined prior distribution. This enables the alignment of learned pulmonary nodule feature distributions between the source domains and generalizes them to the target domain for detection.

Our main contributions are as follows:

- We propose the FASS module to incorporate global features by leveraging the spatial relationship between pulmonary nodules and vascular structures to integrate global features.

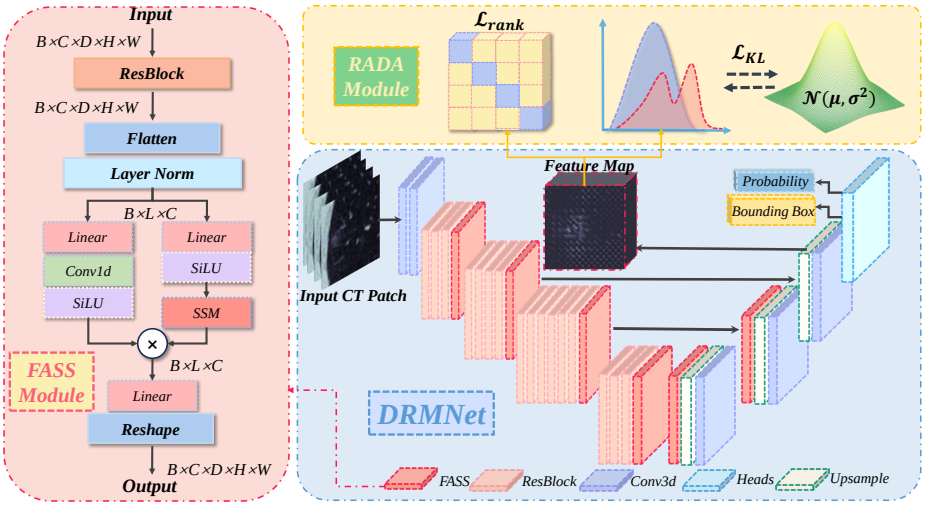

Fig. 2: An overview of the proposed DRMNet. DRMNet is an end-to-end neural network with two key components: the FASS and RADA modules.

- We introduce the RADA module further to align pulmonary nodule features from different source domains, enabling effective generalization to the target domain.
- We construct a domain generalization dataset GPND for pulmonary nodule detection. It contains two private datasets and two public datasets.

2 Methods

2.1 Problem Formulation

Let $\mathcal{X} \in \mathbb{R}^{D\times H\times W\times C}$ represent the input (feature) space, $\mathcal{Y} \in \mathbb{R}$ denote the labels space, $\mathcal{B} \in \mathbb{R}^6$ denote the space of bounding boxes. One *domain* is defined as a joint distribution P_{XY} on $\mathcal{X} \times \mathcal{Y}$. During the training process, there are K datasets $\{S_k\}_{k=1}^{K}$ available, where $S_k = \left\{x_n^k, \{y_{mn}^k, Box_{mn}^k\}_{m=1}^{M_n}\right\}_{n=1}^{N}$ with index n includes input $x_n^k \in \mathcal{X}$ and M_n number of discrete labels $y_{mn}^k \in \mathcal{Y}$ and their corresponding bounding boxes $Box_{mn}^k \in \mathcal{B}$. The goal of the proposed model is to learn a mapping $\mathcal{X} \to (\mathcal{Y} \times \mathcal{B})^U$, that generates U accurate proposals from a single input x_T sampled from an unseen domain.

2.2 Architecture of the DRMNet

As shown in Fig. 2, our proposed DRMNet consists of three parts: the U-shaped backbone with FASS modules to capture information from lung tissues, the

RADA module to minimize KL divergence between source domains, and the heads used to generate the proposals. The loss function of the total model is:

$$\mathcal{L}_{total} = \sum_{n,k} \mathcal{L}_{cls} + \mathcal{L}_{reg} + \mathcal{L}_{rank} + KL(p(\mathcal{Z} \mid \mathcal{X}) \parallel \mathcal{N} \sim (0,1)), \tag{1}$$

where $\mathcal{L}_{cls}$ denotes the Focal Loss and $\mathcal{L}_{reg}$ denotes the Smooth L1 Loss, $\mathcal{L}_{rank}$ is defined by $\partial \mathcal{L}_{rank}/\partial Z = UV^T$, where U and V are obtained from Singular Value Decomposition (SVD). $\mathcal{L}_{rank}$ and $KL(\cdot)$ will be discussed in detail in the RADA module.

2.3 Feature-Augmented State-Space

In pulmonary nodule detection, vessels and bronchioles usually appear as continuous tubular structures on chest CT. Although malignant pulmonary nodules are usually near blood vessels, their imaging features often appear isolated and spheroidal. These subtle differences require identification across multiple slices. Extracting global features mirrors a physician's approach—analyzing multiple slices to distinguish nodules from lung tissues. An optimal model should integrate knowledge from diverse domains and adapt to novel distributions, efficiently encoding, retrieving, and adjusting information for precise decision-making [15]. Existing detectors [14, 21, 22] are fundamentally based on CNNs and thus struggle to store and access the large-scale information needed at each step.

The Mamba is inspired by the continuous system that maps a 1-dimensional function $x_t \to y_t \in \mathbb{R}$ and can be denoted as:

$$\begin{cases} h^{'}(t) = Qh(t) + Rx(t) \\ y(t) = Sh(t) \end{cases}, \tag{2}$$

where the evolution matrix $Q \in \mathbb{R}^{N \times N}$ and projection matrices $R \in \mathbb{R}^{N \times 1}, S \in \mathbb{R}^{1 \times N}$ are its parameters, and $h(t)$ represents a hidden state. The FASS module is proposed to acquire the advantage of U-Net and Mamba for global and local contexts in pulmonary nodule detection. As shown in Fig. 2, each FASS module contains Residual blocks followed by Instance Normalization and Mamba blocks. The Mamba block consists of two streams. The first stream expands features by a linear layer, 1D convolution layer, SiLU function, and the SSM layer. The second branch also expands features by a linear layer and SiLU function. Then, the features from the two streams are integrated and transposed to the original size. Unlike other Mamba works [10, 29], we discard the patch embedding operation, as the detection performance is highly susceptible to the scanning patch order.

2.4 Region-Aware Distribution Alignment

The information extracted from different domains contains a large amount of specific features related to pulmonary nodules. However, due to variations in distribution between these domains and the significant morphological differences of pulmonary nodules, the learned feature representation still suffers from noise

interference. We further introduce a RADA module to regularize the latent feature space among multiple source domains and match the distribution of features to a distribution prior. Inspired by the linear dependency algorithm [11], a feature space z_n^k from one domain can be represented by other domains as $z_n^k = \alpha_1 z_n^1 + \cdots + \alpha_{k-1} z_n^{k-1} + \alpha_{k+1} z_n^{k+1} + \cdots + \alpha_K z_n^K$, where $\{\alpha_1, \alpha_k, \ldots, \alpha_K\}$ denotes the parameters for each feature space, $\alpha_k \geq 0$ and $\|\alpha_k\| \leq M$, M means a Lipschitz constant. We theoretically facilitate the Kullback-Leibler (KL) divergence to demonstrate that when the latent features in the source domain approach a normal distribution, the latent features in the target domain will also tend towards a normal distribution, thereby reducing differences between the source and target domain features. In detail, the KL divergence can be formulated as $KL(p(\mathcal{Z} \mid \mathcal{X}) \parallel \mathcal{N} \sim (0,1))$, where $\mathcal{Z}$ denotes the latent features. The distribution on latent variables can be defined as $p_i(z), i = \{1, 2, \ldots, K, T\}$, the normal prior $\mathcal{N} \sim (0,1)$ as $p_G(z)$, where K, T represent the number of source domain and target domain, respectively. Then, the latent distribution of a sample $x_{i_T}^T$ from the target domain is defined as $p(z_{n_T}^T \mid x_{n_T}^T) = \sum_{k=1}^K \alpha_k p(z_{n_K}^k \mid x_{n_K}^k)$, which follows the normal distribution. By leveraging the $x \geq \log(1+x)$ and $\int p_T(z)dz = \int p_k(z)dz = 1$, we are able to derive an upper bound:

$$
\begin{aligned}
KL(p_T(z) \parallel p_G(z)) &= \sum_{k=1}^{K} \alpha_k \int_z p_k(z) \log \frac{p_T(z)}{p_G(z)} dz \\
&= \sum_{k=1}^{K} \alpha_k \int_z p_k(z) \log \frac{p_k(z)\left[1 + (p_T(z)/p_k(z) - 1)\right]}{p_G(z)} dz \\
&= \sum_{k=1}^{K} \alpha_k \int_z p_k(z) \{\log \frac{p_k(z)}{p_G(z)} + \log\left[1 + (p_T(z)/p_k(z) - 1)\right]\} dz \\
&\leq \sum_{k=1}^{K} \alpha_k \int_z p_k(z) (\log \frac{p_k(z)}{p_G(z)} + p_T(z)/p_k(z) - 1) dz \\
&\leq \sum_{k=1}^{K} \alpha_k \int_z p_k(z) \log \frac{p_k(z)}{p_G(z)} dz \\
&\leq \sum_{k=1}^{K} \alpha_k KL(p_k(z) \parallel p_G(z)).
\end{aligned}
\tag{3}
$$

The equation above indicates that when $KL(p_k(z) \parallel p_G(z))$ is minimized, the $KL(p_T(z) \parallel p_G(z))$ for target domain will also be minimized. Given an input $\mathcal{X} = \left\{x_i^k\right\}$ sampled from the source domain, we can obtain its feature map $\mathcal{Z}$ through the posterior distribution $p(z \mid x)$ parameterized by FASS module. To exploit the intrinsic feature of positive anchors from multiple source domains and reduce the impact of noise on domain-invariant feature alignment, we flatten $\mathcal{Z}$ as 1D Z and design rank regularization $rank(Z)$ by using SVD as $Z = USV^T$, where S is a diagonal matrix containing all the singular values. The previous linear dependency methods [7, 12] are tailored for classification tasks, relying on predefined terms associated with categories to optimize the calculation of singular values. In contrast, our approach, designed for detection tasks, directly

utilizes positive anchor boxes for singular value computation, reducing additional computation.

3 Experiments

3.1 GPND Dataset

The GPND dataset comprises four subsets sourced from public datasets PN9 and LUNA16, along with private datasets PONSD and GGO. The PONSD dataset focuses on postoperative patients, emphasizing small nodes. The GGO dataset includes cases with ground glass opacity (GGO), a high-risk factor often missed on CT scans. These private datasets, sourced from a top-tier Chinese hospital, undergo stringent desensitization and quality checks. The LUNA16 dataset only labeled nodules larger than 3 mm, and PN9 labeled nodules of varying sizes and morphologies, which are multidomain.

3.2 Evaluation Metrics and Implementation Details

The performances of detection systems are evaluated by the Free-Response Receiver Operating Characteristic (FROC) [16], which is employed in follow-up research for pulmonary nodule detection. Detection performance is quantified based on the mean recall achieved at pre-defined false positive rate levels per scan (FP/scan), specifically at 1/8, 1/4, 1/2, 1, 2, 4, and 8 on average. We compute the 95% confidence interval using bootstrapping with 1000 bootstraps. The four subsets are all randomly split into 7:1:2 for training, validation, and testing.

We conduct experiments on an Ubuntu server with 2 Nvidia 4090 GPUs (24G) using CUDA 12.3 and PyTorch 2.1.1. Additionally, we use SANet [14] as the backbone, and we employ a Stochastic Gradient Descent (SGD) to minimize costs. The base learning rate is established at 0.01, along with momentum and weight decay coefficients of 0.9 and 1×10^{-4}, respectively. Training involves 40 total epochs, with the learning rate decreasing to 0.001 after 32 epochs and 0.0001 after 36 epochs.

3.3 Effectiveness of DRMNet

Comparsion with State-of-the-Art DG methods We first compare DRMNet with state-of-the-art DG methods, including LDDG [7], DGER [27], AIDA [6], SMDD [5], and STDR [18]. We conduct four experiments for each method, with one subset of the test set chosen as the target domain and the remaining three subsets of the training set serving as the source domains. The DeepAll refers to the absence of any domain generalization method. The results, presented in Table 1, are obtained by tuning hyperparameters extensively. Notably, all DG methods outperform DeepAll, with DRMNet achieving superior performance by integrating effective feature learning and domain alignment, considering global and local information. Specifically, DRMNet significantly outperforms

Table 1: Domain Generalization results for pulmonary nodule detection (%). Each subset is treated as the target domain in turn, with the remaining three as source domains. Average denotes the mean recall across all four trials.

Method	Target	0.125	0.25	0.5	1	2	4	8	FROC
DeepAll	PN9	5.40	8.00	12.27	19.15	26.56	35.29	46.68	21.91
	GGO	24.89	35.52	51.06	66.69	80.55	89.1	93.65	63.06
	PONSD	1.83	4.15	10.26	24.34	39.41	54.8	64.84	28.52
	LUNA16	44.93	56.48	66.19	72.11	75.67	81.85	84.98	68.89
	Average	19.26	26.04	34.94	45.57	55.55	65.26	72.54	45.59
LDDG [7]	PN9	8.92	12.92	21.98	30.94	37.04	45.99	56.06	30.55
	GGO	39.08	56.57	69.55	77.85	86.63	91.22	95.55	73.78
	PONSD	3.29	6.44	12.61	24.66	37.51	46.97	58.38	27.12
	LUNA16	45.01	57.79	68.92	73.38	77.45	80.72	84.21	69.64
	Average	24.08	33.43	43.26	51.71	59.66	66.22	73.55	50.27
DGER [27]	PN9	6.00	12.01	21.31	30.91	40.07	49.24	57.49	31.00
	GGO	38.82	49.62	63.25	76.75	87.59	92.56	94.32	71.85
	PONSD	1.11	3.23	9.39	18.69	35.58	50.29	64.22	26.07
	LUNA	37.23	49.69	60.43	68.31	73.15	78.26	81.72	64.11
	Average	20.79	28.64	38.59	48.66	59.1	67.58	74.44	48.26
AIDA [6]	PN9	7.18	14.37	23.77	34.79	45.45	53.88	62.3	34.53
	GGO	51.15	58.46	71.43	82.94	90.78	93.85	95.97	77.80
	PONSD	1.37	5.36	14.49	28.82	44.09	55.82	67.69	31.09
	LUNA	46.79	54.71	60.53	67.83	75.48	80.29	85.26	67.27
	Average	26.62	33.22	42.55	53.60	63.95	**70.96**	**77.8**	52.67
SMDD [5]	PN9	8.07	11.75	16.75	26.11	38.88	47.58	54.04	29.02
	GGO	28.96	43.53	58.42	74.55	82.17	88.44	92.84	66.99
	PONSD	1.92	3.85	10.39	23.54	40.70	54.57	65.88	28.69
	LUNA	55.83	61.36	70.04	74.81	79.83	81.56	84.73	72.59
	Average	23.69	30.12	38.90	49.75	60.40	68.04	74.37	49.32
STDR [18]	PN9	13.11	18.14	25.31	32.11	43.05	51.17	56.90	34.26
	GGO	36.77	50.76	64.15	77.24	86.93	89.84	92.52	71.17
	PONSD	1.34	4.83	12.61	22.90	39.54	56.02	67.28	29.22
	LUNA	48.25	56.55	60.85	68.10	74.05	83.40	88.04	68.47
	Average	24.86	32.57	40.73	50.08	60.90	70.11	76.20	50.78
DRMNet	PN9	12.66	19.16	29.16	35.84	42.95	51.22	59.90	35.84
	GGO	68.63	75.27	80.44	87.66	90.87	94.07	95.97	84.70
	PONSD	3.24	9.79	17.9	28.63	42.96	53.61	61.33	31.07
	LUNA	51.96	62.85	69.09	75.24	79.19	82.70	87.91	72.70
	Average	**34.12**	**41.77**	**49.14**	**56.84**	**63.99**	70.40	76.28	**56.08**

other methods in recall when FP/scan is less than 4, an advantage deemed clinically significant. Our findings indicate that all methods generalize better to GGO and LUNA16 but struggle with PN9 and PONSD, likely due to differences in nodule sizes and morphologies between datasets. By contrast, GGO and LUNA16 tend to have larger nodule diameters and more uniform nodule types, reducing source-target domain discrepancies. Qualitative results, depicted in Fig. 3, high-

FASS	LR	KL	Average FROC
			45.59
✓			52.15
✓	✓		53.31
✓		✓	53.10
	✓	✓	49.67
✓	✓	✓	**56.08**

Table 2: Ablation Study on the key components of DRMNet (%).

Methods	Average FROC
VMamba [29]	52.26
UMamba [13]	54.60
SegMamba [20]	53.69
EMNet [1]	53.84
DGMamba [10]	55.15
Ours	**56.08**

Table 3: Ablation Study on different methods based on Mamba (%).

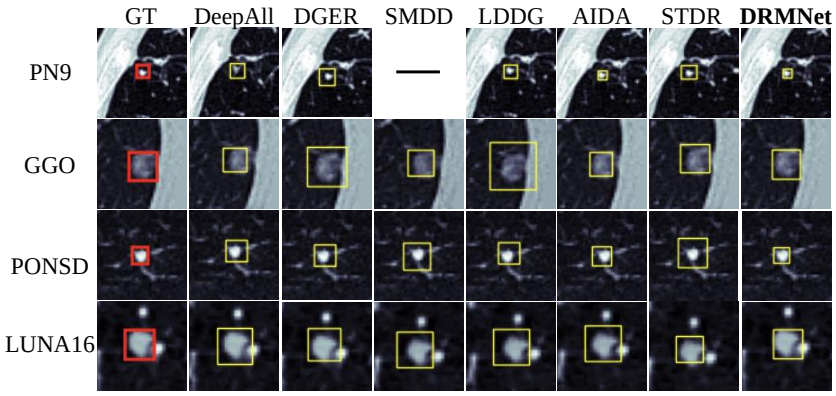

Fig. 3: Visualization results of different domain generalization methods. The first row to the fourth row shows the results on different target domains. "-" indicates the method does not detect the corresponding nodule.

light the ability of our model to generate more accurate predicted bounding boxes compared to others.

Ablation Studies In Table 2, we verify the effectiveness of two modules. Specifically, we conduct separate experiments on the Low-Rank (LR) and KL components of the RADA module. The experimental results show that combining all the modules leads to the best performance. Furthermore, we investigate the influence of the FASS module by removing it from the architecture. The result indicates a noticeable degradation in the FROC performance, underscoring the significance of extracting global information across multiple slices. Overall, combining these components enhances the performance, thus highlighting the importance of both FASS and RADA modules. Furthermore, Table 3 presents the performance of other Mamba-based methods across datasets, highlighting the superiority of our approach.

4 Conclusion

In this paper, we propose a DRMNet to address domain generalization in pulmonary nodule detection. First, we design the FASS module to extract global information across slices, memorizing correlations about pulmonary nodules at the feature level. Next, a RADA module is introduced to reduce discrepancies between domain feature distributions. Extensive experiments on the GPND dataset derived from real clinical data and publicly available benchmark datasets demonstrate the effectiveness and superiority of our proposed method.

Acknowledgments. This work was funded in part by the Key Research and Development Program of the Department of Science and Technology of the Tibet Autonomous Region under Grant XZ202402ZY0003, in part by the National Natural Science Foundation of China under Grant 62172289, and in part by Sichuan University Postdoctoral Science Foundation under Grant 2023SCU12092.

Disclosure of Interests. The authors have no competing interests in the paper.

References

1. Chang, A., Zeng, J., Huang, R., Ni, D.: Em-net: Efficient channel and frequency learning with mamba for 3d medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 266–275. Springer (2024)
2. Chu, X., Jin, Y., Zhu, W., Wang, Y., Wang, X., Zhang, S., Mei, H.: Dna: Domain generalization with diversified neural averaging. In: International Conference on Machine Learning. pp. 4010–4034. PMLR (2022)
3. Dou, Q., Coelho de Castro, D., Kamnitsas, K., Glocker, B.: Domain generalization via model-agnostic learning of semantic features. Advances in neural information processing systems **32** (2019)
4. Gu, A., Dao, T.: Mamba: Linear-time sequence modeling with selective state spaces. arXiv preprint arXiv:2312.00752 (2023)
5. Jiang, H., Gao, M., Li, H., Jin, R., Miao, H., Liu, J.: Multi-learner based deep meta-learning for few-shot medical image classification. IEEE Journal of Biomedical and Health Informatics **27**(1), 17–28 (2022)
6. Kovacs, B., Netzer, N., Baumgartner, M., Eith, C., Bounias, D., Meinzer, C., Jäger, P.F., Zhang, K.S., Floca, R., Schrader, A., et al.: Anatomy-informed data augmentation for enhanced prostate cancer detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 531–540. Springer (2023)
7. Li, H., Wang, Y., Wan, R., Wang, S., Li, T.Q., Kot, A.: Domain generalization for medical imaging classification with linear-dependency regularization. Advances in neural information processing systems **33**, 3118–3129 (2020)
8. Liao, F., Liang, M., Li, Z., Hu, X., Song, S.: Evaluate the malignancy of pulmonary nodules using the 3-d deep leaky noisy-or network. IEEE transactions on neural networks and learning systems **30**(11), 3484–3495 (2019)

9. Liu, X., Thermos, S., O'Neil, A., Tsaftaris, S.A.: Semi-supervised meta-learning with disentanglement for domain-generalised medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24. pp. 307–317. Springer (2021)
10. Long, S., Zhou, Q., Li, X., Lu, X., Ying, C., Luo, Y., Ma, L., Yan, S.: Dg-mamba: Domain generalization via generalized state space model. arXiv preprint arXiv:2404.07794 (2024)
11. Ma, A.J., Yuen, P.C.: Linear dependency modeling for feature fusion. In: 2011 International Conference on Computer Vision. pp. 2041–2048. IEEE (2011)
12. Ma, A.J., Yuen, P.C., Lai, J.H.: Linear dependency modeling for classifier fusion and feature combination. IEEE transactions on pattern analysis and machine intelligence **35**(5), 1135–1148 (2012)
13. Ma, J., Li, F., Wang, B.: U-mamba: Enhancing long-range dependency for biomedical image segmentation. arXiv preprint arXiv:2401.04722 (2024)
14. Mei, J., Cheng, M.M., Xu, G., Wan, L.R., Zhang, H.: Sanet: A slice-aware network for pulmonary nodule detection. IEEE transactions on pattern analysis and machine intelligence **44**(8), 4374–4387 (2021)
15. Mobiny, A., Yuan, P., Cicalese, P.A., Moulik, S.K., Garg, N., Wu, C.C., Wong, K., Wong, S.T., He, T.C., Nguyen, H.V.: Memory-augmented capsule network for adaptable lung nodule classification. IEEE Transactions on Medical Imaging **40**(10), 2869–2879 (2021)
16. Setio, A.A.A., Traverso, A., De Bel, T., Berens, M.S., Van Den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B., et al.: Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. Medical image analysis **42**, 1–13 (2017)
17. Tang, H., Zhang, C., Xie, X.: Nodulenet: Decoupled false positive reduction for pulmonary nodule detection and segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22. pp. 266–274. Springer (2019)
18. Wang, H., Chen, J., Zhang, S., He, Y., Xu, J., Wu, M., He, J., Liao, W., Luo, X.: Dual-reference source-free active domain adaptation for nasopharyngeal carcinoma tumor segmentation across multiple hospitals. IEEE Transactions on Medical Imaging (2024)
19. Wang, J., Lan, C., Liu, C., Ouyang, Y., Qin, T., Lu, W., Chen, Y., Zeng, W., Yu, P.: Generalizing to unseen domains: A survey on domain generalization. IEEE Transactions on Knowledge and Data Engineering (2022)
20. Xing, Z., Ye, T., Yang, Y., Liu, G., Zhu, L.: Segmamba: Long-range sequential modeling mamba for 3d medical image segmentation. arXiv preprint arXiv:2401.13560 (2024)
21. Xu, R., Liu, Z., Luo, Y., Hu, H., Shen, L., Du, B., Kuang, K., Yang, J.: Sgda: Towards 3d universal pulmonary nodule detection via slice grouped domain attention. IEEE/ACM Transactions on Computational Biology and Bioinformatics (2023)
22. Xu, R., Luo, Y., Du, B., Kuang, K., Yang, J.: Lssanet: A long short slice-aware network for pulmonary nodule detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 664–674. Springer (2022)

23. Xu, X., Wang, C., Guo, J., Yang, L., Bai, H., Li, W., Yi, Z.: Deepln: a framework for automatic lung nodule detection using multi-resolution ct screening images. Knowledge-Based Systems **189**, 105128 (2020)
24. Yoon, C., Hamarneh, G., Garbi, R.: Generalizable feature learning in the presence of data bias and domain class imbalance with application to skin lesion classification. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV 22. pp. 365–373. Springer (2019)
25. Zhang, L., Wang, X., Yang, D., Sanford, T., Harmon, S., Turkbey, B., Wood, B.J., Roth, H., Myronenko, A., Xu, D., et al.: Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. IEEE transactions on medical imaging **39**(7), 2531–2540 (2020)
26. Zhang, R., Xu, Q., Yao, J., Zhang, Y., Tian, Q., Wang, Y.: Federated domain generalization with generalization adjustment. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3954–3963 (2023)
27. Zhao, S., Gong, M., Liu, T., Fu, H., Tao, D.: Domain generalization via entropy regularization. Advances in Neural Information Processing Systems **33**, 16096–16107 (2020)
28. Zhou, K., Liu, Z., Qiao, Y., Xiang, T., Loy, C.C.: Domain generalization: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence (2022)
29. Zhu, L., Liao, B., Zhang, Q., Wang, X., Liu, W., Wang, X.: Vision mamba: Efficient visual representation learning with bidirectional state space model. arXiv preprint arXiv:2401.09417 (2024)