Navigating Distribution Shifts in Medical Image Analysis: A Survey

Zixian Su*, Jingwei Guo*, Xi Yang†, Qiufeng Wang, Frans Coenen, and Kaizhu Huang†

***Abstract*—Medical Image Analysis (MedIA) has become indispensable in modern healthcare, enhancing clinical diagnostics and personalized treatment. Despite the remarkable advancements supported by deep learning (DL) technologies, their practical deployment faces challenges due to distribution shifts, where models trained on specific datasets underperform across others from varying hospitals, regions, or patient populations. To navigate this issue, researchers have been actively developing strategies to increase the adaptability and robustness of DL models, enabling their effective use in unfamiliar and diverse environments. This paper systematically reviews approaches that apply DL techniques to MedIA systems affected by distribution shifts. Unlike traditional categorizations based on technical specifications, our approach is grounded in the real-world operational constraints faced by healthcare institutions. Specifically, we categorize the existing body of work into Joint Training, Federated Learning, Fine-tuning, and Domain Generalization, with each method tailored to distinct scenarios caused by Data Accessibility, Privacy Concerns, and Collaborative Protocols. This perspective equips researchers with a nuanced understanding of how DL can be strategically deployed to address distribution shifts in MedIA, ensuring diverse and robust medical applications. By delving deeper into these topics, we highlight potential pathways for future research that not only address existing limitations but also push the boundaries of deployable MedIA technologies.**

***Index Terms*—Medical Image Analysis, Deep Learning, Distribution Shifts, Transfer Learning, Trustworthy AI.**

I. Introduction

MEDICAL image analysis (MedIA) [1] has become a cornerstone of modern healthcare, playing a critical role in enhancing diagnostics [2]–[4], patient monitoring [5], and treatment planning [6]. With the advent of high-resolution imaging technologies and the increasing complexity of medical data, the application of advanced computational tools has become indispensable. Deep learning (DL) technologies [7]–[10], in particular, have revolutionized MedIA by enabling automated and accurate analyses of medical images [11], [12]. These technologies leverage large datasets to train models that can recognize patterns with a precision often surpassing human capabilities [13]. The integration of DL in MedIA not only speeds up diagnostic processes but also offers the potential for personalized healthcare through more accurate patient-specific assessments.

However, the application of deep learning techniques in MedIA faces substantial challenges, primarily due to distribution shifts. These shifts occur because the training data (known as source data) used to develop DL models often come from highly controlled environments or specific populations. When deployed in varied medical settings – like different hospitals, population regions, and time periods – these models encounter data that differ significantly in aspects such as imaging modalities [14], scanning protocols [15], patient populations [16], and temporal changes [17]. These variations expose the models to novel, out-of-distribution patterns (referred to as target data) that they have not been trained to recognize, impairing their ability to generalize effectively; this compromised performance in turn undermines the reliability and effectiveness of DL-based diagnostics. Therefore, addressing these distribution shifts is crucial for the effect and reliable deployment of DL technologies in diverse medical environments.

To this end, this survey focuses on investigating DL-based MedIA under the challenges posed by distribution shifts. In recent years, the research community has actively developed strategies to enhance the adaptability and robustness of DL models. These strategies aim to mitigate the impact of data distribution shifts across diverse medical settings [18], [19]. In real-world healthcare, the successful deployment of DL technologies often encounters various operational constraints that directly leads to different data distribution shift scenarios. These constraints typically stem from several key factors:

- **Data Accessibility:** This aspect concerns the availability of comprehensive datasets for training DL models. The breadth and quality of accessible data impacts how well a model can be trained to handle varied medical conditions, determining the difficulty level of managing the potential data distribution shifts.
- **Privacy Concerns:** Given the sensitive nature of medical data, privacy concerns [20] revolve around the protection of patient information. These considerations often limit the sharing of medical data among different healthcare institutions, creating data silos that exacerbate the potential data distribution shifts.
- **Collaborative Protocols:** Collaboration among healthcare institutions enables collective efforts to improve diagnostic models across diverse settings. By adhering to different protocols, various collaborative methods [21], [22] have been developed while meeting specific require-

Z. Su is with University of Liverpool, Liverpool, UK and Xi'an Jiaotong-Liverpool University, Suzhou, China (E-mail: zixian.su@liverpool.ac.uk).

J. Guo is with Alibaba Group, Beijing, China (E-mail: jingweiguo19@outlook.com).

X. Yang and Q. Wang are with Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu, China (Email: xi.yang01@xjtlu.edu.cn; qiufeng.wang@xjtlu.edu.cn).

F. Coenen is with University of Liverpool, Liverpool, UK (Email: coenen@liverpool.ac.uk).

K. Huang is with Duke Kunshan University, Kunshan, Jiangsu, China (Email: kaizhu.huang@dukekunshan.edu.cn).

*Equal contribution.

†Corresponding authors.

ments to alleviate the potential distribution shifts.

Building on these practical considerations while deploying DL models, we categorize existing efforts to manage distribution shifts in MedIA into a hierarchy from simple to hard:

- **Joint Training:** This approach is feasible when both the source and target data are accessible and there is no privacy concerns. This scenario often occurs when multiple health institutions agree to share their own data, facilitating joint model training [23], [24] and thereby enhancing model adaptability across diverse settings.
- **Federated Learning:** When multiple institutions seek to cooperate without exposing their distinct datasets due to privacy concerns, federated learning [25] offers a powerful solution. It enables collaborative model improvements across different institutions by training models locally on each dataset and aggregating the learned models without centralizing data storage.
- **Fine-tuning:** When synchronous collaborations are not allowed for addressing data distribution shifts with privacy concerns, fine-tuning [26], [27] emerges as an effective remedy. This involves using a well pre-trained model and then fine-tuning it on new datasets to transfer learned knowledge to unfamiliar domains.
- **Domain Generalization:** When data from unseen domains that require model adaptation is inaccessible or unknown, training a model that is generalizable enough to withstand distribution shifts is essential [28], [29]. This involves preparing for unforeseen challenges by developing models that can generalize from the data currently available for training to any potential new environments.

In this survey, we present a nuanced understanding of how deep learning can be strategically deployed to address distribution shifts in MedIA, facilitating the development of diverse and robust applications. While a few surveys have also explored the impact of distribution shifts on MedIA and summarized solutions, our work stands apart in several critical ways. For instance, [18] primarily focuses on domain adaptation (DA) within MedIA, categorizing existing methods based on the degree of DL model supervision while [19] emphasizes Domain Generalization (DG) and organizes existing methods according to common MedIA workflows. Although DA and DG are significant topics with profound impacts on MedIA, these surveys [18], [19] concentrate on the technical aspects of existing approaches, treating MedIA primarily as an application domain. Their classifications, rooted in the intricacies of DL techniques, often overlook the real-world medical constraints that give rise to different distribution shift scenarios. Consequently, they fail to provide a detailed, step-by-step guide addressing the impact of medical data variations.

Unlike them, our approach is grounded in the practical, operational constraints faced by healthcare institutions, examining current DL techniques in light of the real factors affecting MedIA under distribution shifts. Moreover, while some surveys have explored these issues within medical contexts, they often restrict their discussions to specific scenarios (e.g., heart/lung/brain), such as [30]–[33], lacking a comprehensive exploration of distribution shifts within MedIA. Our survey addresses these gaps by offering direct and actionable strategies for deploying DL models under the unique operational constraints encountered in real-world applications. This not only serves as a practical guide for medical professionals on employing deep learning to tackle genuine medical challenges but also underscores the transformative potential of DL technologies in MedIA. By highlighting operational constraints and providing tailored solutions, our survey deepens the understanding and broadens the application of deep learning in MeIA, thereby enhancing both the field and the integration of artificial intelligence in healthcare.

TABLE I
DEFINITION OF MATHEMATICAL NOTATIONS.

Notation	Definition	Notation	Definition
$\mathcal{X}, \mathcal{Z}, \mathcal{Y}$	Input, feature, output space	$P(\cdot)$	Prior Knowledge
$\mathbf{x}, \mathbf{z}, y$	Input, feature, label variables	$\mathcal{L}(\cdot,\cdot)$	Loss function
$\mathcal{D}_t$	Target Domain	$q_\theta(\cdot)$	Predictive function
$\mathcal{D}_s$	Source Domain	$\mathcal{M}(\cdot)$	Manipulation function
$p(\cdot)$	Probability distribution	$f(\cdot)$	Feature mapping function

A. Problem Definition

In this section, we formalize the distribution shift problem with notations defined in Table I for easy reading. A domain $\mathcal{D}$ is a joint distribution $p(x, y)$ defined on the input-output space $\mathcal{X} \times \mathcal{Y}$, where random variables $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ denote the input data and the output label, respectively. We typically deal with two distinct datasets, known as the **source** and **target** domains. The **Source Domain** $\mathcal{D}_s = \{(x, y) \sim p_s(x, y)\}$. comprises medical images x such as X-rays or MRI scans, each paired with a label y that might be categorical information regarding disease diagnosis or the segmentation mask. The **Target Domain** $\mathcal{D}_t = \{(x, y) \sim p_t(x, y)\}$ originates from a different but related distribution to that of the source. For instance, they might come from different medical imaging devices or patient populations. Note that for both source and target distributions, $p_s(x, y) = p_s(x)p_s(y|x)$ and $p_t(x, y) = p_t(x)p_t(y|x)$. We take the standard covariate shift assumption as **Distribution Shift**, i.e., $p_s(y|x) = p_t(y|x)$ and $p_s(x) \neq p_t(x)$. In this situation, the model $q_\theta(y|x)$ solely trained on the source domain cannot well represent the true, domain-invariant distribution $p(y|x)$. Therefore, a variety of research concentrates on adjusting $q_\theta(y|x)$ to maximize its predictive performance on the target distribution.

II. BACKGROUND

A. Distribution Shifts in Medical Image Analysis

The efficacy of DL models largely hinges on the assumption that the training and testing data are independently and identically distributed (i.i.d). However, this assumption often does not hold in the complex and diverse environment of clinical practice. The inherent heterogeneity in medical imaging, arising from different modalities, varying protocols, diverse patient demographics, and temporal shifts, introduces significant distribution shifts. In the following, we provide a concise illustration (see visual examples in Fig. 1):

- Imaging Modalities: Medical imaging encompasses a range of modalities, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), X-rays, and Ultrasound, each producing images with unique characteristics. A model trained on data from one modality might not generalize well to another, given the inherent differences in image textures, contrasts, and anatomical representations.
- Scanning Protocols: Even within the same modality, images can vary based on the imaging protocols and equipment used. Factors such as magnetic field strength in MRI, radiation dose in CT, and ultrasound machine settings can introduce significant variations in the images.
- Patient Demographics: Differences in patient populations, such as age, gender, and ethnicity, as well as variations in disease manifestations, can lead to substantial differences in imaging data. For instance, pediatric images are markedly different from adult images.
- Temporal Shifts: Longitudinal studies and data collected over extended periods often encounter shifts due to the progression of diseases, the impact of treatments, and changes in physiological states. As a result, models trained on historical data may not perform optimally on current or future data. These shifts challenge the performance of models when applied outside their training domain, a common occurrence given the variety of imaging techniques used in healthcare.

III. Categorization and Frameworks

This section provides an overview of our categorization rationale and framework for addressing MedIA under distribution shifts. We primarily consider the real-world operational constraints encountered when deploying DL techniques in MedIA, such as Data Accessibility, Privacy Concerns, and Collaborative Protocols. These factors shape different scenarios under distribution shift in MedIA, leading to the classification of exisiting DL techniques into four main categories: Joint Training, Federated Learning, Fine-tuning, and Domain Generalization. A visual depiction of this classification is provided in Fig. 2. Within each major category, we further subdivide the techniques into prominent subdomains in the field, primarily based on differences in label rate and diversity of data. These subdomains are then ranked according to their learning difficulty (see Table II). Furthermore, to help researchers quickly identify the technical types of methods they are interested in, we categorize the methods within each subdomain under three aspects as elaborated in the following:

- **Data Management**: Focuses on increasing the model's exposure to varied data scenarios through strategic data augmentation, selection, and translation techniques.
- **Model Design**: Involves modifying the structural and strategic elements of frameworks to enhance adaptability and robustness to changes in data distribution.
- **Optimization Strategy**: Encompasses advancements in training process adjustments, optimizing how models learn from data exhibiting distribution shifts.

It is important to note that although methods are classified by their primary innovation, many integrate multiple strategies

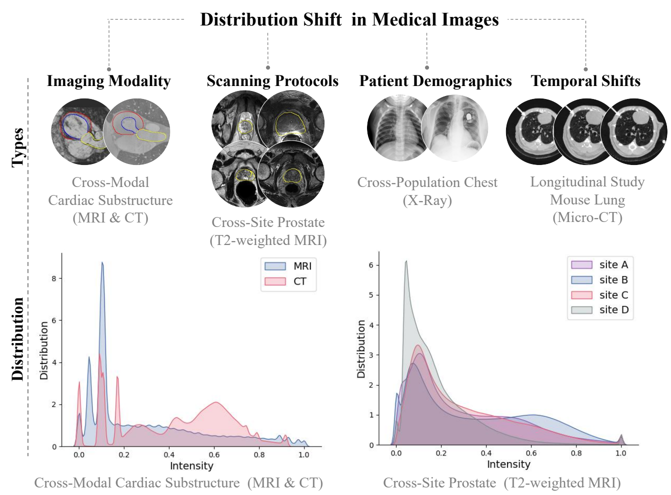

Fig. 1. Illustration of medical imaging distribution shifts, showcasing from Imaging Modalities (Cardiac Substructure [14]), Scanning Protocols (Cross-Site Prostate [15]), Patient Demographics (Cross-Population Chest [34], [35]), and Temporal Shifts (Mouse Lung [36]).

across these aspects. This structured taxonomy not only aids researchers in identifying methods suited to specific problems but also facilitates a comparative and systematic analysis.

IV. Joint Training

Joint Training is a crucial domain adaptation strategy in MedIA, particularly effective when target data is freely accessible and privacy concerns are minimal. This method excels in environments where healthcare institutions can collaboratively share data, creating the ideal conditions for joint model training. Such collaboration significantly enhances the adaptability of models across varied medical settings by integrating both source and target data. Typically, the source dataset is fully-labeled, whereas the target dataset often exhibits varying labeling rates due to changes in medical scenarios, introducing complexities to DL model training. In response, a variety of joint training strategies have emerged, each designed to address the specific challenges posed by fluctuating label availability on target data. These methods are categorized based on the level of target supervision, ranging from Supervised to Semi-supervised, and Unsupervised Joint Training.

A. Supervised Joint Training

Supervised Joint Training is a domain adaptation strategy where models are concurrently trained on both the source and target domain data, leveraging labeled data from both to enhance performance despite domain shifts. This method is particularly valuable when the target domain has significantly less labeled data than the source, as relying solely on target data would yield inadequate model performance. In Supervised Joint Training, the strategy involves integrating different modalities or varying views of data, and often includes synthesizing data to mitigate the data shortage problem in the target domain. These methods effectively utilize the structural and distributional characteristics of data from both domains,

TABLE II
TAXONOMY OF DL TECHNIQUES DESIGNED FOR MEDIA UNDER DISTRIBUTION SHIFTS.

Methods	Settings	Medical Scenarios	Difficulty	Pros	Cons
Joint Training	Supervised	Analysis of multimodal medical imaging (e.g., using both PET and CT images simultaneously).	Low	Utilizes different but complementary information to enhance robustness	Need to synchronize distinct data sources
	Semi-supervised	Target labeled data is scarce. By combining a small amount of labeled data with a large amount of unlabeled data, models can be trained to improve accuracy, especially for rare diseases where obtaining sufficient labeled samples is challenging.	Medium	Uses available labels efficiently	Highly dependent on the quality of labeled data
	Unsupervised	The target hospital or device lacks sufficient labeled data. For example, adapting a model to a new hospital or new equipment dataset without the additional labeled data from the target environment.	Medium	No need for labeled target data	May struggle with very distinct domain shifts
Federated Learning	Supervised	Collaborative research across multiple hospitals. This allows hospitals to improve the model collaboratively while maintaining patient data privacy.	Medium	Preserves privacy, good for multi-institutional data	High communication overhead and complexity
	Semi-supervised	Same with the Supervised setting for privacy preserving while the label is scarce for the involved hospitals.	High	Leverages the existing unlabeled data efficiently, practical for real-world collaborations	Achieving convergence can be difficult and slow
Fine-tuning	Supervised	For smaller hospitals with limited resources, they can utilize models pre-trained on large datasets and finetune them with a small amount of local data to quickly deploy effective diagnostic tools.	Low	Quickly adapts pre-trained models to new tasks with minimal data	Risk of overfitting, especially with limited data
	Unsupervised	Similar to the above but source data are unavailable due to privacy or security concerns.	High	Ideal for strict privacy settings, adapts using only target domain data	May suffer from model degradation and sensitive to supervision signal
Domain Generalization	Multi-source	Multiple source domains (e.g., from different hospitals/modalities/scanners) with labels to improve model ability to generalize to new, unseen medical environments.	Medium	Strong generalization	Requires diverse source data
	Single-source	Data diversity is limited, e.g., source dataset comes from only one type of MRI scanner.	High	Simple implementation	Limited effectiveness across wider domain shifts

making full use of all available labeled data to effectively bridge the gap between the source and target domains.

1) Data Management

Cross-modal Translation. Cross-modal translation plays a pivotal role in addressing the challenge of integrating data from diverse imaging modalities, which often exhibit distinct intensity and texture characteristics. This technique facilitates the conversion of data between modalities, such as from MRI to CT images, enabling the use of a unified dataset for training despite the inherent discrepancies. By synthesizing data from one modality in the form that resembles another, cross-modal translation helps to overcome the shortage data problem and enhances the robustness of the training process. Specifically, Generative Adversarial Networks have proven to be particularly effective for cross-modal translation [37]–[40] by creating high-quality synthetic images that maintain the domain-specific characteristics of the target modality. For example, the shape-consistency approach [37] leverages GANs for volume-to-volume translation, ensuring that the structural integrity of medical images is preserved across modalities.

2) Model Design

Architecture Variations. Novel architectural designs are crucial for addressing domain adaptation challenges. For instance, the domain-adaptive two-stream U-Net, applied for electron microscopy image segmentation [41], features a dual-stream architecture that supports selective weight sharing between source and target domains. This design enhances adaptability by allowing the model to fine-tune its responses to the unique characteristics of each domain. Similarly, the Multi-Site Network (MS-Net) for cross-site prostate segmentation [23] incorporates Domain-Specific Batch Normalization (DSBN). DSBN effectively manages inter-site variability by providing distinct feature normalization for each site, ensuring that the model remains robust across diverse MRI datasets.

3) Optimization Strategy

Metric Learning. Metric Learning has proven instrumental in maintaining high generalization performance across different data domains. A notable application is demonstrated in [42], where metric learning is employed to enhance domain adaptation for Wireless Capsule Endoscopy. This approach utilizes triplet loss, a form of contrastive learning, which effectively minimizes the distance between embeddings of samples with the same labels from different domains while maximizing the distance between samples with different labels from the same domains. By doing so, it ensures that the model can accurately interpret and classify medical images regardless of the specific device version.

B. Semi-supervised Joint Training

Semi-supervised Joint Training, also referred to as Semi-Supervised Domain Adaptation (SSDA), is a cutting-edge machine learning strategy aimed at transferring knowledge from a well-labeled source domain to a target domain with scarce labels. This approach is vital in situations where acquiring comprehensive labels for the target domain is impractical due to cost or time constraints. The primary challenges of SSDA include maximizing the utility of limited labeled data and a larger volume of unlabeled data in the target domain, as well as mitigating distribution discrepancies between the domains.

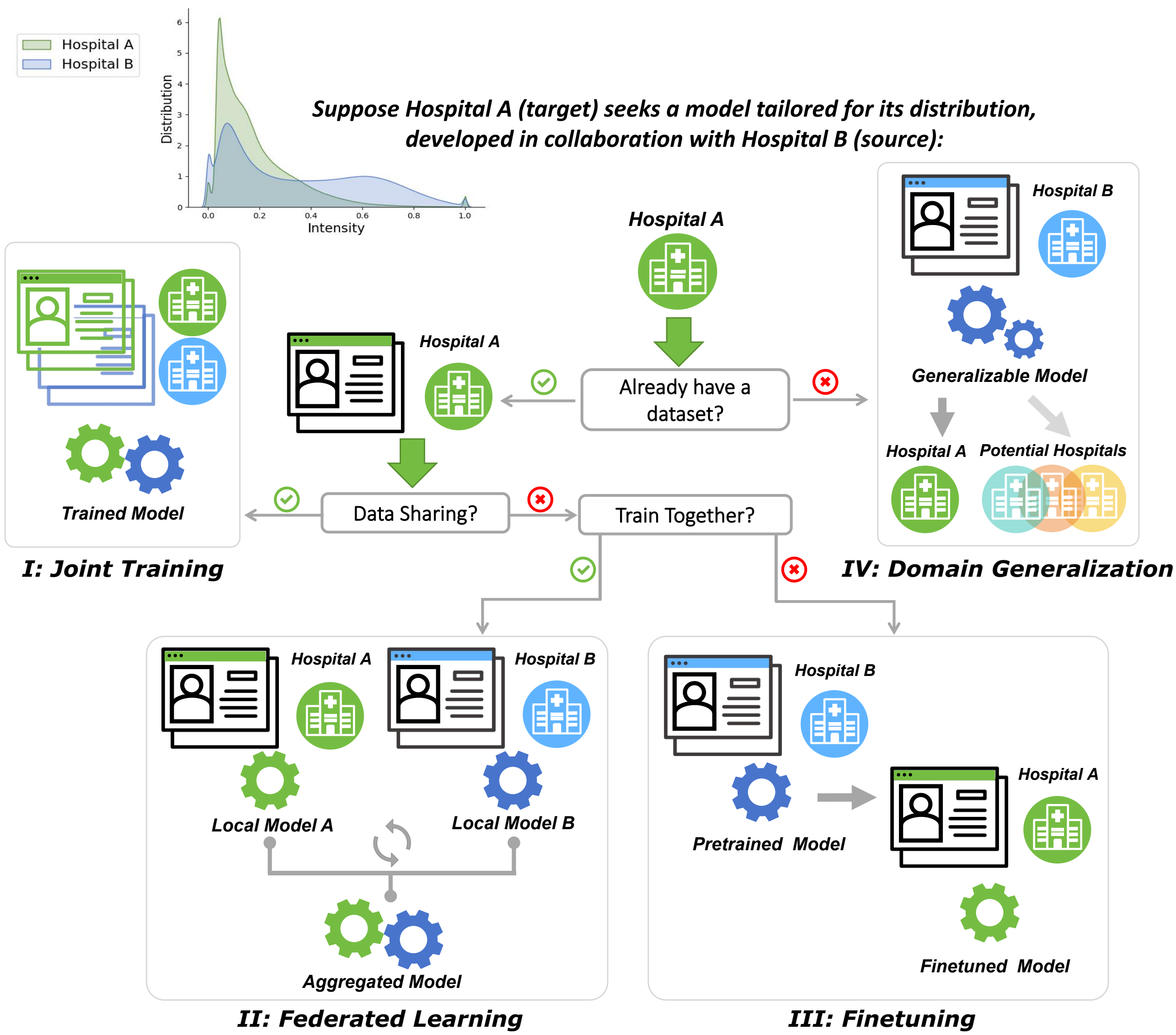

Fig. 2. The diagram categorizes existing deep learning techniques into four main approaches, each addressing real-world operational constraints including Data Accessibility, Privacy Concerns, and Collaborative Protocols. In Joint Training, hospitals collaborate by sharing data for model training. Federated Learning enables collaboration without direct data sharing, maintaining privacy. Fine-tuning adapts pre-trained models from one institution to another. Domain Generalization develops models that generalize across diverse settings, even without access to target data, to mitigate distribution shifts.

1) Data Management

Pseudo-labeling. Pseudo-labeling is a powerful technique in semi-supervised joint training that leverages large volumes of unlabeled data to enhance model training. This method involves generating artificial labels for unlabeled data based on the most confident predictions of the model, thereby expanding the training dataset effectively. Specifically, it employs the learned features from training data to generate pseudo-labels for unlabeled data. In domain adaptation, its effectiveness hinges on prioritizing model's predictions exceeding a defined threshold that are more likely to be accurate:

$$\hat{y}_u = \begin{cases} \arg\max q_\theta(x_u), & \text{if } \max q_\theta(x_u) > \tau \\ \text{ignore}, & \text{otherwise} \end{cases} \tag{1}$$

where x_u represents an unlabeled sample from the target domain, $q_\theta(x_u)$ denotes the model's probabilistic predictions, $\hat{y}_u$ is the pseudo-label assigned to x_u, and τ is a threshold defining the confidence level above which the labels are considered reliable. The pseudo-label loss is computed as:

$$\mathcal{L}_u = \sum_{x_u \in \mathcal{X}_u} \mathbb{I}\left(\max q_\theta(x_u) > \tau\right) \cdot \mathcal{L}(\hat{y}_u, q_\theta(x_u)) \tag{2}$$

where $\mathbb{I}(\cdot)$ is an indicator function that selects high-confidence samples. The total training objective combines the loss from labeled data and high-confidence pseudo-labeled data:

$$\mathcal{L}_{total} = \mathcal{L}_l(\mathcal{Y}_l, q_\theta(\mathcal{X}_l)) + \beta \mathcal{L}_u(\hat{\mathcal{Y}}_u, q_\theta(\mathcal{X}_u)) \tag{3}$$

where β is a balancing factor between the source and pseudo-labeled losses.

To mitigate the risk of error propagation, which can occur if incorrect labels are used for training, enhancements are made to ensure the quality of these pseudo-labels. For example, [43] employs transformation-invariant, highly-confident predictions in the target dataset for self-training purposes, ensuring that the model is less likely to learn from noisy, less reliable labels. Meanwhile, [44] enhances the robustness of pseudo-labeling by calculating the variance between the original image and its Fourier-transformed counterpart, providing a more stable basis for generating reliable pseudo-labels. These strategies significantly improve the utility of pseudo-labeling, making it a vital tool for utilizing unlabeled data in domain adaptation.

2) Model Design

Self-ensembling. Self-ensembling is an advanced learning strategy that effectively exploits both labeled and unlabeled data with the consistency between models. This method trains multiple versions of a model, each subjected to distinct input perturbations, and employs consistency regularization to ensure uniform predictions across these variations. Typically, this technique under SSDA setting is implemented via a "teacher-student" model, where a stable, pre-trained "teacher" model q_θ^{tc} guides a less-trained "student" model q_θ^{st}. The overall training process is governed by two primary loss functions:

$$\begin{aligned}\mathcal{L}_{total} &= \mathcal{L}_l(\mathcal{Y}_l, q_\theta^{tc}(\mathcal{X}_l)) \\ &+ \lambda\mathcal{L}_{cons}(q_\theta^{tc}(\mathcal{X}_u)), q_\theta^{st}(\mathcal{X}_u^{aug}))\end{aligned} \tag{4}$$

where supervised loss $\mathcal{L}_l$ is for measuring discrepancies in the teacher's predictions on labeled input $q_\theta^{tc}(\mathcal{X}_l)$, and consistency loss $\mathcal{L}_{\text{cons}}$ is for aligning the teacher's predictions on unlabeled input $q_\theta^{tc}(\mathcal{X}_u)$ with the student's on perturbed inputs $q_\theta^{st}(\mathcal{X}_u^{aug})$. The two loss functions are balanced by a regularization parameter λ.

Following this framework, [45], [46] introduce this techniques into achieveing semi-supervised domain adaptation. The core innovation of these frameworks lies in the strategic employment of dual-teacher models: one teacher model enhances intra-domain knowledge through self-ensembling techniques, while the other facilitates inter-domain knowledge transfer using image translation models such as CycleGAN [47]. This approach leverages the consistency of model outputs across different views of the same data, enhancing the model's ability to generalize across diverse scenarios.

Adversarial Learning. Adversarial leaning is fundamentally embodied by Generative Adversarial Networks (GANs). It can be conceptualized as a game between two players:

$$\begin{aligned}\min_G \max_D V(D, G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] \\ + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))]\end{aligned} \tag{5}$$

where generator (G) aims to produce data that is indistinguishable from real data by transforming input noise z, sampled from a noise distribution $p_z(z)$, and discriminator (D) aims to correctly classify real data x and generated data $G(z)$. Real data x is sampled from the true data distribution $p_{\text{data}}(x)$. This interaction forms a min-max game where the generator seeks to deceive the discriminator into accepting its outputs as real, while the discriminator improves at identifying the differences between real and generated data. Through this adversarial process, G refines its outputs to reduce discrepancies, indirectly generating domain-invariant features for domain adaptation. As one notable method, COVID-DA [48] is designed to distinguish between closely related conditions such as pneumonia and COVID-19, particularly when labeled data is scarce. This method uses a unique classifier separation scheme along with an adversarial network to overcome the task difference and domain discrepancy simultaneously.

Novel Training Strategies. [49] explores the richness of multi-modal data through a novel asymmetric co-training approach. By segmenting the learning process into two distinct components that each addresses specific aspects of domain adaptation and semi-supervised learning task, this strategy avoids the domination of the source domain data thus facilitates more effective domain adaptation.

3) Optimization Strategy

Metric Learning. Metric Learning within the context of semi-supervised joint training is distinctly innovative. In this scenario, [50] adopts a metric learning strategy characterized by a disentangled paradigm. This approach separates style and content into distinct embedding spaces. Such separation facilitates independent contrastive learning for each aspect, allowing the model to adapt more effectively to variations in data distributions.

C. Unsupervised Joint Training

Unsupervised Joint Training, commonly known as Unsupervised Domain Adaptation (UDA), is an advanced machine learning framework that facilitates the transfer of knowledge from a richly-labeled source domain to a completely unlabeled target domain. The central challenge of UDA lies in the absence of labels in the target domain, necessitating techniques that can align the underlying data distributions of both domains to enable accurate predictions on the target dataset. Key strategies include domain invariant feature extraction and distribution alignment.

1) Data Management

Cross-Modal Translation. Cross-modal translation, employing techniques such as Generative Adversarial Networks (GANs) and frequency-based methods, is pivotal in transforming how we address domain differences by converting data from the source domain S to closely resemble the target domain T. This transformation is formalized as:

$$\tilde{\mathcal{X}} = G(\mathcal{X}_s; \theta_G) \quad \text{where} \quad G : \mathcal{D}_s \to \mathcal{D}_t \tag{6}$$

Here, G represents a generative model that minimizes domain discrepancies to align source domain $\mathcal{D}_s$ with target domain $\mathcal{D}_t$. The adaptation's effectiveness hinges on reducing the domain discrepancy metric d, which measures differences between the adapted $\tilde{\mathcal{X}}$ and target $\mathcal{X}_t$:

$$\min_{\theta_G} d(\tilde{\mathcal{X}}, \mathcal{X}_t) = \min_{\theta_G} d(G(\mathcal{X}_s; \theta_G), \mathcal{X}_t) \tag{7}$$

Once the data $\tilde{\mathcal{X}}$ closely resembles $\mathcal{X}_t$, it can be used to train models with source domain labels $\mathcal{Y}_s$, significantly enhancing the model's ability to generalize across domains and mitigate image scarcity in specialized fields.

The application of GANs has evolved significantly, beginning with methods like the pixel-to-pixel (pix2pix) translation and advancing to more complex implementations. For instance, [51] utilizes an end-to-end unsupervised method for enhancing contrast in cataract fundus image based on pix2pix framework. Then, MADGAN [52] breaks the constraints of paired images, contributing to anomaly detection in complex brain structures. SASAN [53] takes a further step that incorporating self-attention modules in its GANs, enhancing focus on specific anatomical details during image translation. Subsequently, the utilization of CycleGANs [47] marks another significant advancement, enabling unpaired image translations for cross-domain chest X-ray disease recognition [54], [55] and hip joint bone segmentation [56]. This translation process is further refined in [57] with dual-scheme (source-target/target-source) fusion and [58] with attention mechanism. Integrating disentangled representations into GAN frameworks, as seen in [59]–[64], significantly advances domain adaptation by separating content from style, enhancing adaptation efficiency. Complementing the adversarial nature of GANs, frequency-based methods [65], [66] introduce a novel perspective. They assume that the style information is stored in low frequency components and high frequency components represents more structural information, and thus translate the images by replacing the low frequency components. Finally, techniques like singular value decomposition for noise adaptation in retinal OCT images [67] highlight the innovation and adaptability in this field, which is tailored to specific imaging modalities or diagnostic requirements.

Pseudo-labeling. In unsupervised joint training, as all target data labels are unknown, it becomes more challenging to make accurate pseudo-label predictions using traditional techniques. Research has since advanced the pseudo-labeling concept by integrating pseudo-labeling and adversarial learning to enhance the process [68]. Subsequent studies have built on this foundation, each offering unique improvements to address issues such as noisy labels [69] and enhancing label reliability through methods like iterative self-training [70], contrastive learning [71], and entropy constraints [72]. The specialized applications of pseudo-labeling are further explored in studies [73]–[75]. For example, [73] focuses on nuclei instance segmentation and classification, utilizing pseudo-labels derived from prototype features. [74] breaks new ground in cell detection with a pseudo-cell-position heatmaps. [75] innovates by incorporating pseudo-labeling into tagged-to-cine MRI synthesis task, employing a Bayesian uncertainty mask for selective pseudo-label generation.

2) Model Design

Adversarial Learning. Adversarial Learning is widely used for the implicit alignment between domains at feature or/and pixel level due to the absence of target labels. At the feature level, techniques such as the plug-and-play adversarial domain adaptation network (PnP-AdaNet) [76] aligns features across different scales for segmentation tasks. Similarly, [77] aligns extracted contents for cross-modality segmentation. Other studies focus on prediction space alignment at the pixel level for various medical imaging tasks [78]–[81]. Integrated approaches that apply adversarial training at both feature and output levels are explored in studies like [24], [82]–[87]. Innovations in this field also include enhanced discriminators and local discriminators that focus on specific region alignment, introducing spatial-aware and class-specific attentions to refine the adversarial loss and improve model adaptability across domains [88]–[91].

Self-ensembling. Initial studies by [92] applied self-ensembling to gray matter MRI segmentation. Subsequent applications include breast MRI segmentation [93] and pose estimation in operating rooms [94]. More advanced techniques combine adversarial training and self-ensembling for addressing domain shifts in cross-institutional gliomas studies [95], optic disc and cup segmentation [96] as well as cardiac substructure segmentation [97]. Other significant developments include MT-UDA [98], which introduces a multi-teacher framework, and [99] further integrates frequency and spatial domain through multi-teacher distillation. Moreover, [100] explores a 'student-to-partner' paradigm during various training stages.

Graph-based Methods. Graph-based methods are increasingly utilized in cross-domain medical image analysis due to their capability to capture complex spatial structures and relationships. This approach models image elements – ranging from individual pixels to entire regions – as nodes in a graph, with edges formed based on criteria like spatial proximity and similarity in intensity or texture. The core of this method involves a graph $G = (V, E)$, with V representing the vertices and E the edges, which are weighted according to the mentioned criteria. This setup facilitates the use of graph convolutional networks (GCNs) [101]–[103], which leverage the graph structure for learning, described mathematically as:

$$H^{(l+1)} = \sigma\left(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} H^{(l)} W^{(l)}\right) \tag{8}$$

where $H^{(l)}$ represents the features at each node in layer l, A, D is adjacency and degree matrix, $W^{(l)}$ is the weight matrix for layer l, σ denotes the activation function. This process effectively leverages node features and graph topology for a comprehensive analysis. Applications of this method include feature disentanglement [104] for domain-invariant learning [105] with GCN, graph Laplacian decomposition for brain imaging alignment across domains [106], attention-guided GCN for identifying major depressive disorder [107], and a class-aware GCN classifier with domain-specific features for predicting lymph node metastasis in gastric cancer [108]. Other notable implementations like [109] extends beyond traditional methods by incorporating an online sub-graph scheme, [110] employs GCNs with a meta-learning strategy targeting at small-sized pancreatic cancer features. Studies like [111], [112] focus on enhancing feature alignment and understanding inter-category relationships using graph-based techniques.

3) Optimization Strategy

Statistical Discrepancies Minimization. Quantifying and subsequently minimizing the statistical discrepancies between source and target domain feature spaces serves as a key approach. This paradigm, rooted in the hypothesis that reducing such discrepancies aids in model adaptation, predominantly employs measures like Kullback-Leibler (KL) divergence, Maximum Mean Discrepancy (MMD), and Correlation Alignment (CORAL). For example, the adaptation with MMD [113] between source and target domains can be mathematically formulated as:

$$\mathcal{L}(D_s, D_t) = \frac{1}{n_s^2}\sum_{i=1}^{n_s}\sum_{j=1}^{n_s} k(z_s^i, z_s^j) + \frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t} k(z_t^i, z_t^j) - \frac{2}{n_s n_t}\sum_{i=1}^{n_s}\sum_{j=1}^{n_t} k(z_s^i, z_t^j) \tag{9}$$

where $k(z, z') = \exp\left(-\frac{\|\mathrm{vec}(z)-\mathrm{vec}(z')\|^2}{2b}\right)$ is the Gaussian kernel function defined on the vectorization of tensors z and z' with bandwidth parameter b, z_s, z_t are the multi-layer fused feature of the source and target domains. This distance assess how similar or dissimilar the feature representations of the two domains are, and the goal is to adjust the feature representations such that the distance between the empirical distributions (as represented by the kernel functions) is minimized, thereby enabling adaptation.

In the medical adaptation field, [114] explores the utility of MMD for domain adaptation in breast and thyroid lesions in ultrasound images. [115] leverages KL divergence to synchronize the prior distribution of the synthesized and the real target distribution. [116] estimates the mutual information with KL divergence between the reconstruction output and segmentation result, so as to benefit each other. [117] enforces recursively conditional Gaussian (RCG) as the joint distribution prior, inheriting the closed form of the KL divergence term in the variational objective to make large-sacle tasks computationally tractable. [118] uses MultiKernel Maximum Mean Discrepancy (MK-MMD) in aligning feature distributions in breast ultrasound images. Beyond these widely adopted metrics, novel metrics have been developed to suit specific medical tasks. [119] proposes the Characteristic Function (CF) Distance, transforming feature distributions to frequency domain for discrepancy calculations. [120] introduces Domain Sanity Loss, focusing on anatomical features like centroid distance and plausibility in vertebrae prediction.

V. Federated Learning

Federated Learning is a pivotal model training approach designed to handle data heterogeneity while preserving the privacy of each client. It is particularly valuable in MedIA for alleviating data distribution shifts, allowing for collaborative enhancements across multiple healthcare institutions without the need to centralize their data. This decentralized method ensures the privacy of patient information, making it a practical solution for scenarios where medical data cannot be openly shared. One prominent example of Federated Learning in practice is FedAvg [25], which forms the basis for many modern implementations. In this model, each participating institution trains a local model on its own data, thereby maintaining the confidentiality of sensitive information. These institutions then send their model updates – commonly in the form of weights or gradients – to a central server. The server aggregates these updates to enhance the global model, which is shared back with all participating institutions after a few iterations. The mathematical formulation is detailed as follows:

$$\text{Local Update: } \theta_k^{(t+1)} = \theta_k^{(t)} - \eta\nabla L_k(\theta_k^{(t)}) \tag{10}$$

$$\text{Global Update: } \theta^{(t+1)} = \sum_{k=1}^{K}\frac{n_k}{n}\theta_k^{(t)} \tag{11}$$

where θ_k and θ respectively represent the parameters of the local model for the k-th client and the global model. Each client k contributes n_k data points, which together total n data points across K clients. The learning rate is denoted by η, and $\nabla L_k(\theta_k^{(t)})$ refers to the gradient of the loss L_k with respect to the local model parameters at the k-th client.

Similar to the Joint Training category, Federated Learning methods can also be classified based on the degree of data labeling. However, unlike Joint Training, which primarily focuses on the label availability of target data, Federated Learning treats both source and target data as clients that play similar roles. Each client trains a local model that contributes to the quality of the expected global model. Therefore, in this section, our primary concern regarding label availability extends to all clients. Based on this, Federated Learning for MedIA under distribution shifts can be divided into Supervised and Semi-supervised Federated Learning. Besides, it is worth noting that, given that medical settings often feature well-characterized source datasets, the need for Unsupervised Federated Learning approaches [121], [122] is generally minimal.

A. Superivsed Federated Learning

Early research in the medical field utilizing the FedAvg [25] algorithm targeted a wide range of medical tasks, from Brain Tumor Segmentation [123] to the detection of COVID-19 lung abnormalities [124], [125], MRI Reconstruction [126], Diabetic Retinopathy Classification of OCT Data [127], and Breast Cancer Histopathological Image Classification [128]. These initial applications laid the foundation for using Federated Learning to process sensitive medical data across distributed datasets while maintaining privacy. Further studies explored the influence of factors like the number of healthcare providers, dataset size, communication strategies, and architecture types on FL performance in medical contexts [129], [130]. As the field progressed, benchmarks were established to assess the effectiveness of various FL algorithms in managing data heterogeneity [131] across diverse medical datasets. Moving beyond the basic FedAvg paradigm, current advancements have focused on addressing issues like Data Heterogeneity and Client Drift [132], [133], which arise from non-IID data distributions among clients. These shifts can significantly affect model performance, prompting researchers to develop strategies for the accuracy of FL models in healthcare settings.

1) Data Management

Data Augmentation. In Federated Learning for MedIA under distribution shifts, cross-client data augmentation plays a crucial role in managing the inherent diversity and imbalance of data across different clients. This strategy is designed to enhance the uniformity of feature representations across participating clients, thus improving the overall robustness and accuracy of the federated model. Techniques such as Fourier transform-based methods [134] are particularly effective, as they allow for the sharing and interpolation of frequency domain information among clients, promoting a more consistent feature representation across varied datasets. Specifically, HarmoFL [135] leverages frequency information to unify amplitude components across clients, which aids in maintaining consistent low-level visual features. Other systematic augmentation techniques [136]–[139] explore various augmentation strategies to combat data diversity and imbalance. These techniques vary in their approaches but collectively contribute to a more equitable and effective training process, enhancing the ability of federated models to generalize across diverse environments and data conditions.

2) Model Design

Novel Architecture Design. Some strategies specifically address distribution shifts by developing tailored model architectures. For instance, SU-Net [140] enhances standard U-Net with inception modules and dense blocks to manage multi-scale challenges effectively. Similarly, FedDAvT [141] leverages Transformer architecture to facilitate domain adaptation for Alzheimer's disease diagnostics. Adversarial and generative networks are introduced to refine federated learning, focusing on aligning or adapting feature spaces across different clients [142], [143].

3) Optimization Strategy

Metric Learning. Several methods utilize metric learning to enhance consistency between different clients in federated settings. For instance, FedIIC [144] implements two-level contrastive learning to optimize both intra- and inter-client feature consistency, ensuring uniformity in the learned representations. FedCL [145] focuses on reducing the feature distance between successive local and global models, which helps stabilize the training process. Similarly, FedDP [146] improves model uniformity across clients by penalizing inconsistencies during the learning phase. Additionally, LC-Fed [147] employs contrastive site embedding and makes prediction-level adjustments to enhance personalization.

Novel Training Strategies. Novel Training Strategies are being explored to enhance the efficacy and adaptability of models. For example, FedSM [148] optimizes model selection based on data distribution during inference, while FedCross [149] employs a unique approach that involves sequential training without the need for model aggregation. Additionally, strategies such as the Dropout, Mixture of Experts and Split Learning have been introduced to improve model effectiveness [150]–[153]. These innovative methods collectively contribute to more secure and resilient model training and deployment.

Aggregation Weight Calibration. In Federated Learning for MedIA, aggregation weight calibration is a sophisticated optimization strategy that refines how global model updates are weighted, taking into account more than data volume. This method involves adjusting the influence of each client's local update on the global model by considering factors such as the stage of training, client performance, and similarity between client models and the global model. For example, [154], [155] highlight strategies where weights are calibrated based on the training progress and the performance metrics of clients. Additionally, the similarity-based approach [156]–[159] assesses how closely aligned each client's data distribution or model parameters are with the global model. This alignment influences their weights during aggregation, promoting updates that are more representative of the overall data characteristics. Moreover, FedAWA [160] introduces an innovative twist by employing reinforcement learning to dynamically adjust client weights. This system continually learns and updates based on data distribution and feedback from client performance, optimizing the aggregation process to ensure the global model remains robust and accurate across varying conditions.

Parameter Calibration. Parameter calibration also plays a crucial role, specifically for addressing the conflict between the local and global models. It involves strategically adjusting model parameters to ensure that the collective learning benefits all participating clients. Efforts include rescaling local parameters [161] and mixing local and global gradients [162] to enhance model convergence and stability. [163] proposes a Deputy-Enhanced Transfer strategy at the client site. It firstly leverages a deputy model to receive aggregated parameters from the server, and then smoothly transfers the global knowledge to the personalized local model. Some other strategies emphasize fairness, such as those aiming to equalize training loss by adjusting the model parameters such that all hospitals have a similar training loss [164]. This approach ensures that no single client's data disproportionately influences the model, thus maintaining uniformity in model performance regardless of the data source.

B. Semi-supervised Federated Learning

In the diverse landscape of Federated Learning, Semi-Supervised Federated Learning (SSFL) emerges as a pivotal area of exploration, particularly suited to complex environments like healthcare, where only a subset of clients possess fully labeled data, while a significant portion operates with unlabeled datasets. By incorporating techniques from semi-supervised learning, SSFL effectively utilizes sparse labels to extrapolate knowledge and enhance learning from the extensive unlabeled data available. This approach not only broadens the applicability of Federated Learning in the medical field but also adeptly addresses the latent data heterogeneity challenges that emerge when the lack of clear labels obscures underlying data variations.

1) Data Management

Pseudo-labeling. Several innovative approaches have been developed to enhance the utility of pseudo-labeling in federated settings. [165] introduces a novel method that integrates prototype-based pseudo-labeling with contrastive learning, a technique also employed by [166]. Additionally, [167] enhances pseudo-label generation by incorporating a self-supervised rotation loss, which provides consistent regularization across unlabeled datasets. Further, [168] improves the connection between labeled and unlabeled data by aligning disease relationships across clients, effectively compensating for the lack of task-specific knowledge in unlabeled clients and enhancing the extraction of discriminative information from unlabeled samples.

2) Model Design

Transformer-based Architecture. Transformer offers a robust framework for leveraging both labeled and unlabeled data within a single client. For example, [169] exemplifies a specialized approach where a self-supervised learning framework is implemented using Transformer architectures. This method starts with masked image modeling, a self-supervised task that trains the model to predict the portions of images that are intentionally obscured. This phase harnesses the abundant unlabeled data, allowing the model to learn rich, generalized features without requiring too many explicit labels.

3) Optimization Strategy

Advanced Optimization Strategies in SSFL also address the dual challenges of data scarcity and distribution heterogeneity. One innovative approach is the Federated Drift Mitigation (FedDM) framework [170], which achieves robust gradient aggregation by resolving conflicts between gradients at different network layers, as guided by the historical gradients of the global model. Another strategic implementation is FedCy [171], designed for surgical phase recognition. This method integrates dual training objectives: it applies consistency learning to exploit the temporal and spatial consistencies in the unlabeled data, alongside contrastive learning techniques to enrich the learning from sparsely labeled data.

VI. Fine-tuning

Fine-tuning plays a vital role in enhancing the adaptability and performance of pre-trained models across a wide range of applications. This process involves adjusting a model that has been pre-trained on a large, generic (source) dataset to perform effectively on a different, often smaller and more specialized (target) dataset. In medical scenarios, Fine-tuning proves particularly effective when privacy concerns preclude open data sharing, and synchronous collaborations among different healthcare institutions are impractical or excessively costly. This strategy enables medical institutions to leverage pre-existing models and adapt them with minimal data exchange, effectively addressing privacy and collaboration constraints in MedIA. Based on the availability of labeled data on the target domain, fine-tuning methods are classified into supervised and unsupervised approaches. As we move from supervised to unsupervised settings, the complexity increases but so does the significance of the application, offering broader adaptability to real-world challenges where labeled data are limited.

A. Supervised Fine-tuning

Supervised Fine-tuning stands out as a potent method for enhancing diagnostic accuracy in MedIA. This technique primarily involves applying specific pre-trained networks, such as VGG [172] and AlexNet [173], initially trained on general images like ImageNet [174], to more specialized medical imaging tasks. Research exemplified by studies [175]–[178] demonstrates how these models transition to applications in medical imaging, including tumor classification and chest X-ray analysis, leveraging their capability to generalize features across diverse visual domains for precise medical diagnostics. Fine-tuning these networks often requires minimal adaptation design, making it a straightforward approach to boost performance in medical tasks. Notable successes also include adapting networks for Alzheimer's diagnosis [26] and employing the Med3D network for detailed lung segmentation and nodule classification [179].

1) Model Design

Novel Strategies and Structures. Beyond simply evaluating on different pre-trained network architectures, some research have focused on novel strategies and structures for rapid and accurate domain adaptation, while preserving existing knowledge. [180] introduces ContextNets, a memory-augmented network for seamless domain adaptation in semantic segmentation without the need for extensive retraining. In contrast, [181] employs Elastic Weight Consolidation to maintain performance by encoding information from previous tasks, without extra data storage. Furthermore, [182] optimizes batch normalization to swiftly adjust to new domains while maintaining shared convolutional layers across all domains.

B. Unsupervised Fine-tuning

Unsupervised Fine-tuning in MedIA is an innovative response to the constraints of traditional supervised fine-tuning that rely heavily on labeled target datasets which are often unavailable in healthcare scenarios. This approach, crucial in healthcare where rapid adaptation is required to varying patient data, is characterized by two primary branches: Source-Free Domain Adaptation (SFDA) [183] and Test-time Adaptation (TTA) [184]. Both are designed to adapt models dynamically to new and changing conditions without the need for source data at the time of inference, thus directly addressing the challenges of data privacy. SFDA achieves this by transferring knowledge learned during training and applying it to new test samples through adaptive modules or auxiliary self-supervised tasks, such as rotation prediction. This allows the model to train on the target distribution for multiple epochs before making predictions, providing a proactive adaptation approach. On the other hand, TTA takes on a more challenging task by requiring the model to adapt in real-time to a continuous stream of test data, making no prior adjustments during the training phase. This method is model-agnostic and focuses on immediate, on-the-fly adjustments to effectively process and respond to incoming data. Both strategies share the common goal of enabling efficient model adaptation in unsupervised settings, ensuring that medical diagnostics remain robust and accurate even when faced with data that significantly deviates from previously seen examples.

1) Data Management

Pseudo-labeling. An intuitive solution for SFDA/TTA is to use the source model to generate pseudo labels [185]–[188] for the target domain data and thus convert the problem into a supervised one on the target domain. However, these pseudo labels often contain noise due to domain discrepancies, making it essential to refine them for accuracy. Techniques include adaptive pseudo-labeling which uses dual-classifiers to enhance label confidence in [189], denoised pseudo-labeling with uncertainty and prototype distance estimation for precise segmentation [190], and employing shape compactness metrics for label reweighting [191]. Additionally, in [192], a system integrates an image quality assessor and an irregular structure detector is developed to select optimal pseudo-labels for training. [193] uses the greatest union mask of multiple predictions to generate proxy labels for model fine-tuning, while [188] selects low-entropy pixels as reliable labels and applies contrastive learning to tighten the target feature distribution.

Image Generation. Image generation techniques facilitate the adaptation of models to new domains by enriching the dataset with varied and representative examples. For example, [194] utilizes basic image augmentation combined with causal interventions to generate diverse datasets that ensure consistent predictions and the elimination of confounding factors. Similarly, [195] employs patch-wise processing augmented with a Transformer structure to enhance data variability effectively, while [196] proposes the first learnable test-time augmentation policy that dynamically selects most effective augmentation techniques. This adaptability allows for optimal model performance even under varying operational conditions. Moreover, some strategies focus on transforming the style of data between the source and target domains to better align the characteristics of the target data with the learned source domain model. For instance, [197] applies autoencoders to adjust test images to resemble source-domain images more closely, enhancing the model's applicability to new data. Additionally, [198] and [191] explore generative techniques. The former uses a class-conditional generative adversarial network to create target-style data from random noise, while the latter leverages Fourier transformation to generate source-like images through a style-mining generator. [199] further innovates by learning a domain-aware prompt that modifies target inputs to better match the source domain style, facilitating smoother domain adaptation.

2) Model Design

Batch Normalization. Batch Normalization has been widely explored in adaptation tasks as normalization statistics are associated with the domain distribution. They can be directly obtained through pre-trained model and taken as the source information. Given a mini-batch $\mathcal{B} = \{x_n\}_{n=1}^{N}$ where $x_n \in \mathbb{R}^F$ is a feature vector (with F denoting the number of feature channels and N the batch size), BN normalizes each feature dimension f as follows:

$$\hat{x}_{n,f} = \frac{x_{n,f} - \mu_{\mathcal{B},f}}{\sqrt{\sigma_{\mathcal{B},f} + \epsilon}} \cdot \gamma_f + \beta_f \tag{12}$$

where $\mu_{\mathcal{B},f}$ and $\sigma_{\mathcal{B},f}$ are the running mean and variance for the f-th feature of mini-batch $\mathcal{B}$, respectively. The parameters γ_f and β_f are the learned scale and shift factors for affine transformation, with ϵ being a small-offset to avoid division by zero. [200] proposes an exponential decay scheme for the normalization statistics in adaptation stage to gradually learn the target domain-specific mean and variance. [199] aligns the source and target normalization statistic discrepancy for learning a prompt to make the target inputs be treated as the source. More recently, [201] explores domain-specific and shareable batch normalization statistics for adaptive BN-based adaptation, while [202] proposes to incorporate the concept of class diversity to address more realistic mini-batch problem.

Novel Strategies and Structures. The field has seen several structural innovations aimed at overcoming specific adaptation challenges. [203] introduces an auxiliary rotation classifier to improve adaptation via self-training. Similarly, [204] utilizes multiple diverse classifiers to address test label distribution shifts, and [205] employs decoder duplication during the adaptation stage to ensemble diverse target inputs. Y-shaped architectures with dual decoders are used for enhanced denoising and segmentation [206], [207]. [208] further develops a supplementary network to adaptively combined with the main outputs during inference.

3) Optimization Strategy

Entropy Minimization. Entropy minimization is widely-used to handle unlabeled data. Mathematically, the entropy of a prediction can be expressed as follows:

$$H(p) = -\sum_{i=1}^{C} p_i \log p_i \tag{13}$$

where p represents the predicted probability distribution over C classes, and p_i is the probability of the i-th class predicted by the model. The goal is to minimize this entropy $H(p)$ across the dataset, thereby encouraging the model to produce more decisive outputs. This approach is first introduced by Tent [184] into general TTA tasks, which proposes minimizing the mean entropy over the test batch to update the affine parameters of the batch normalization layers in the pre-trained model. This strategy has been adopted in many cases [200], [208]–[211] in medical TTAs.

Dynamic Adjustment of Learning Rates. Dynamic adjustment of learning rates based on distribution shifts helps models adapt more effectively during test-training stages. For example, [27] proposes that samples with larger distribution shift should result in larger update. It makes adjustment by calculating the divergence between the model outputs and its nearest neighbors in a memory bank. [212] refines this strategy by assessing category-wise discrepancies with an uncertainty estimation module.

Anatomical Information. Leveraging the anatomical information as a prior for loss design offers a promising direction for enhancing the accuracy and reliability. For instance, [213] utilizes a shape dictionary, integrating general semantic shapes extracted from source data. [214] incorporates the shape information with the signed distance field which measures the distance between any pixel to the nearest object boundary and the relative position. [210], [211] leverages the class-ratio as a supervision which is estimated from anatomical knowledge available in the clinical literature. [215] uses anatomically-derived loss functions that penalizes unrealistic bone lengths and joint angles in 3D pose estimation. [216] proposes contour regularization loss for constraining the continuity and connectivity. [217] expands the lesion click (i.e., the center of the nodule) into an ellipsoid mask, and use it as the supervised information for test training.

VII. Domain Generalization

Domain Generalization (DG) is an advanced deep learning technique designed to prepare models for handling unseen, out-of-distribution data. This challenge is especially relevant in MedIA, where real-world operational constraints often make target datasets from new domains inaccessible or unknown. In such cases, DG methods become essential, enabling models to generalize from available source data to new environments without prior exposure to specific target data. This proactive approach ensures that medical models remain robust and accurate, ready to cope with potential unfamiliar environments. Furthermore, unlike other categories that are typically divided from simple to complex based on the label availability of target data, DG assumes that no target data is available. The complexity of tasks within this category primarily hinges on the nature of the source data. Specifically, DG techniques can be divided into two main types: Multi-source Domain Generalization (MDG) and Single-source Domain Generalization (SDG). MDG capitalizes on the diversity of multiple source datasets to extract and decouple domain-invariant and domain-specific features, thereby enhancing the model's generalizability using the domain-invariant component. Conversely, SDG, limited to a single source, faces greater challenges and often relies on additional data augmentation strategies to increase the model's generalization capabilities under more restrictive conditions.

A. Multi-source Domain Generalization

Multi-source Domain Generalization (MDG) operates under the premise that the unseen target domain shares the commonalities with the source dataset. The main challenge here is effectively extracting and balancing domain-invariant features – which apply across all datasets – and domain-specific features – which are unique to each dataset. Techniques such as feature disentanglement and meta-learning are often employed to address these challenges, helping to enhance the model's ability to generalize while reducing the risk of overfitting to any single source domain.

1) Model Design

Meta-learning. Meta-learning [218] is a powerful strategy for enhancing model generalization across unknown data distributions. This approach involves simulating domain shifts during training through "episodes", where data from multiple sources is split into meta-train $\mathcal{D}_{\text{train}}$ and meta-test $\mathcal{D}_{\text{test}}$ sets. This split mirrors real-world domain shifts, preparing the model for new domains or distributions. The model first learns from the meta-train set and is then tested on the meta-test set to evaluate its adaptability to new situations. Adjustments are made based on its performance to enhance its generalization capabilities. This process is formulated as:

$$\phi^* = \mathit{MetaLearn}(\mathcal{D}_{\text{train}}), \theta^* = \mathit{Learn}(\mathcal{D}_{\text{test}}; \phi^*) \tag{14}$$

where ϕ^* denotes the meta-learned parameters, which are then used to learn the task-specific model parameters θ^* on the meta-test set. Following this framework, [15] introduces a shape-aware meta-learning scheme that incorporates anatomical integrity, [28] combines meta-learning with style-feature flow generation for confounding factors elimination, and [219] uses style-transferred images as meta-tests, designing a new boundary-oriented objective for meta-optimization considering the specific challenges in medical image segmentation.

2) Optimization Strategy

Shape-based Regularization. Shape-based Regularization is a powerful tool, harnessing the continuous and coherent nature of anatomical structures and the domain-invariant characteristics of their contours. Except for combined with the meta-learning approaches [15], [28], [219] for supervision in meta-test optimization, some methods directly use the anatomical knowledge as prior information during training. For example, [220] integrates fixed Sobel kernels for contour enhancement and a convolutional autoencoder for learning anatomical priors, which inversely projects the mask and prediction to the feature space for further alignment.

Latent Space Regularization. Latent Space Regularization focuses on modeling inter-domain relationships and perform regularization in the latent feature space to promote generalization. Notably, [221] introduces a rank regularization term to constrain the complexity of feature representations and restrict the latent features to follow a pre-defined prior distribution, while [222] implements semantic feature regularization during the meta-test phase with dual losses that maintain global inter-class relationships and tighten intra-class features.

B. Single-source Domain Generalization

Single-source Domain Generalization (SDG) presents a unique set of challenges as it relies on data from only one source to prepare models for unseen domains. This restriction is particularly pronounced in the medical field, where variability in data can be extreme and the stakes of accurate generalization are high. The primary challenge in SDG is the limited diversity, which can make models prone to biases and over-fitting, reducing their ability to perform well on novel, out-of-distribution medical data. To combat this, SDG strategies often incorporate robust data augmentation techniques – such as synthetic image generation, geometric transformations, and intensity variations – to artificially expand the dataset's diversity and simulate potential unseen scenarios. Additionally, regularization techniques and invariant feature learning are used to further enhance the model's generalization capabilities.

1) Data Management

Pixel-level Augmentation. Pixel-level Augmentation techniques directly manipulate the pixel values. This method is primarily based on the premise that variations in imaging modalities, acquisition protocols, and hardware can induce significant discrepancies in image characteristics such as texture, intensity, and contrast. For example, [223] introduced BigAug, a deep stacked general transformation approach to systematically evaluate augmentation effects on model generalization. Specialized approaches within the medical field, such as the use of Bézier Curves by [29] to address gray-scale discrepancies, and the simulation of MRI distortions by [224], focus on medical-specific image traits. [225] uses causal inference methods to reflect acquisition shifts and [226] explores category-level augmentation based on class-level representation invariance. [227] combines the augmentation strategies both in [223] and [226]. [228] expands the style space through adversarial training and finds the worst-case style composition to generate the samples. [229] further refines this strategy by introducing randomness to the generated domain through a adaptive instance normalization block, so that the changes are limited to the textures.

Feature-level Augmentation. Some augmentation strategies delves deeper into the model's internal workings, focusing on the manipulation of learned feature representations. For instance, [230] masks features channel-wisely and spatially to generate diverse challenging samples.

VIII. Future Directions

Advancing medical image analysis (MedIA) in the face of distribution shifts requires not only addressing current challenges but also exploring innovative research pathways that can extend the capabilities of deployable technologies. In this section, we identify and discuss several promising directions—namely, continuous learning systems, the utilization of vision-foundation models, and multi-task/multi-modal learning under distribution shifts. By delving into these areas, we aim to highlight potential strategies that can overcome existing limitations and propel the field toward more robust and adaptable MedIA solutions.

A. Continuous Learning Systems for MedIA

In medical imaging analysis, continuous learning systems [231], [232] are essential for adapting to dynamic environments where data distributions evolve over time. As medical practices advance and new clinical data becomes available, models must accommodate these shifts without the need for complete retraining. For instance, data collected over different time periods – such as imaging acquired from new devices, updated imaging protocols, or the emergence of previously unseen disease variants – leads to temporal distribution shifts that challenge conventional static models. Continuous learning addresses these shifts by incrementally updating the model as new data arrives, ensuring that it remains relevant to evolving clinical scenarios. Importantly, these systems are designed to combat catastrophic forgetting, where a model loses performance on previously learned tasks as it incorporates new knowledge. Techniques such as rehearsal, regularization, and memory-based strategies allow continuous learning models to maintain stable performance across a broad spectrum of conditions, ultimately improving adaptability in clinical workflows where the nature of data is in constant flux.

B. Harnessing Vision-Foundation Models for MedIA

The utilization of Vision-Foundation Models (VFMs) presents a promising direction for mitigating distribution shifts in MedIA. For instance, existing research has led to the development of MedSAM [10] and MedCLIP [233], which are fine-tuned versions of SAM [234] and CLIP [235], respectively. MedSAM excels in universal segmentation, allowing it to adapt to diverse medical datasets and imaging protocols, thereby showing its robustness against variations in data distribution. MedCLIP leverages zero-shot learning capabilities, aiming to recognize and analyze unseen medical images based on the embeddings of created prompts for each disease class, which is invaluable in clinical settings with limited labeled data. Further investigation could be the combination of these models for innovative solutions, such as a comprehensive diagnostic tool that integrates segmentation and contextual analysis, providing real-time insights that adjust to new imaging conditions and patient histories. Moreover, leveraging existing large models to enhance the functionality of another model, thereby reducing the impact of distribution shifts, is a promising direction worth exploring. For example, utilizing CLIP to generate visually descriptive sentences related to the segmentation target could enable SAM to effectively perform zero-shot medical segmentation [236]. Future research should explore these synergies, potentially leading to multi-task learning frameworks that simultaneously address various medical tasks, thus enhancing diagnostic accuracy and personalizing treatment strategies. By pursuing these avenues, we can unlock the full potential of VFMs to create resilient and adaptable systems that significantly improve healthcare outcomes.

C. Multi-task/modal Learning under Distribution Shifts

Multi-task [237] and multi-modal [238] learning frameworks offer a promising avenue for addressing the complexities inherent in medical image analysis, particularly when data arises from diverse sources or modalities. These models, designed to jointly process data such as medical images, textual reports, and real-time procedural videos, facilitate a more comprehensive understanding of patient conditions. For example, a multi-modal system could simultaneously perform tumor segmentation from MRI, extract pertinent clinical information from patient records, and analyze surgical video footage to assess tissue responses. Such integrative models are robust against distributional shifts that occur between different data types, ensuring that the system can handle variable quality and types of inputs. Moreover, by leveraging shared representations across tasks and modalities, the model is better equipped to capture general, robust features rather than task-specific nuances, which might overfit to a single type of data.

IX. Conclusion

In this paper, we provide a comprehensive examination of how DL models can be adapted to handle the significant challenge of distribution shifts in MedIA. By categorizing adaptation strategies into Joint Training, Federated Learning, Fine-tuning, and Domain Generalization, we align these methodologies with the practical constraints of Data Accessibility, Privacy Concerns, and Collaborative Protocols that healthcare institutions face. Each strategy offers tailored solutions to specific challenges, ensuring DL models' reliability and effectiveness across various medical environments. Looking forward, refining these adaptive techniques to meet emerging data challenges and advancing technological capabilities will be crucial. Our survey aims to serve as a foundational guide for further research and practical implementation, fostering developments that enhance both the precision and accessibility of MedIA applications in improving patient care.

References

[1] J. S. Duncan and N. Ayache, "Medical image analysis: Progress over two decades and the challenges ahead," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 85–106, 2000.

[2] T. Liu, E. Siegel, and D. Shen, "Deep learning and medical image analysis for covid-19 diagnosis and prediction," *Annual Review of Biomedical Engineering*, vol. 24, no. 1, pp. 179–201, 2022.

[3] V. D. P. Jasti, A. S. Zamani, K. Arumugam, M. Naved, H. Pallathadka, F. Sammy, A. Raghuvanshi, and K. Kaliyaperumal, "Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis," *Security and Communication Networks*, vol. 2022, no. 1, p. 1918379, 2022.

[4] J. Guo, Y. Yang, Q. Wu, J. Su, and F. Ma, "Adaptive active contour model based automatic tongue image segmentation," in *2016 9th International Congress on image and signal processing, BioMedical engineering and informatics (CISP-BMEI)*. IEEE, 2016, pp. 1386–1390.

[5] O. Gozes, M. Frid-Adar, H. Greenspan, P. D. Browning, H. Zhang, W. Ji, A. Bernheim, and E. Siegel, "Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis," *arXiv preprint arXiv:2003.05037*, 2020.

[6] D. Jin, D. Guo, T.-Y. Ho, A. P. Harrison, J. Xiao, C.-K. Tseng, and L. Lu, "Deeptarget: Gross tumor and clinical target volume segmentation in esophageal cancer radiotherapy," *Medical Image Analysis*, vol. 68, p. 101909, 2021.

[7] K. Huang, A. Hussain, Q.-F. Wang, and R. Zhang, *Deep learning: fundamentals, theory and applications*. Springer, 2019, vol. 2.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.

[9] O. N. Manzari, H. Ahmadabadi, H. Kashiani, S. B. Shokouhi, and A. Ayatollahi, "Medvit: a robust vision transformer for generalized medical image classification," *Computers in Biology and Medicine*, vol. 157, p. 106791, 2023.

[10] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images," *Nature Communications*, vol. 15, no. 1, p. 654, 2024.

[11] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, no. 1, pp. 221–248, 2017.

[12] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical image analysis using convolutional neural networks: a review," *Journal of Medical Systems*, vol. 42, pp. 1–13, 2018.

[13] H. Malik, M. S. Farooq, A. Khelifi, A. Abid, J. N. Qureshi, and M. Hussain, "A comparison of transfer learning performance versus health experts in disease diagnosis from medical imaging," *IEEE Access*, vol. 8, pp. 139 367–139 386, 2020.

[14] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of mri," *Medical Image Analysis*, vol. 31, pp. 77–87, 2016.

[15] Q. Liu, Q. Dou, and P.-A. Heng, "Shape-aware meta-learning for generalizing prostate mri segmentation to unseen domains," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2020, pp. 475–485.

[16] J. Fu, S. Bendazzoli, Ö. Smedby, and R. Moreno, "Unsupervised domain adaptation for pediatric brain tumor segmentation," *arXiv preprint arXiv:2406.16848*, 2024.

[17] J. Yang, T. Vetterli, P. P. Balte, R. G. Barr, A. F. Laine, and E. D. Angelini, "Unsupervised domain adaption with adversarial learning (udaa) for emphysema subtyping on cardiac ct scans: The mesa study," in *IEEE International Symposium on Biomedical Imaging*. IEEE, 2019, pp. 289–293.

[18] H. Guan and M. Liu, "Domain adaptation for medical image analysis: a survey," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 3, pp. 1173–1185, 2021.

[19] J. S. Yoon, K. Oh, Y. Shin, M. A. Mazurowski, and H.-I. Suk, "Domain generalization for medical image analysis: A survey," *arXiv preprint arXiv:2310.08598*, 2023.

[20] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, no. 6, pp. 305–311, 2020.

[21] M. Lenga, H. Schulz, and A. Saalbach, "Continual learning for domain adaptation in chest x-ray classification," in *Medical Imaging with Deep Learning*. PMLR, 2020, pp. 413–423.

[22] D. Ng, X. Lan, M. M.-S. Yao, W. P. Chan, and M. Feng, "Federated learning: a collaborative effort to achieve better medical imaging models for individual sites that have small labelled datasets," *Quantitative Imaging in Medicine and Surgery*, vol. 11, no. 2, p. 852, 2021.

[23] Q. Liu, Q. Dou, L. Yu, and P. A. Heng, "Ms-net: multi-site network for improving prostate segmentation with heterogeneous mri data," *IEEE Transactions on Medical Imaging*, vol. 39, no. 9, pp. 2713–2724, 2020.

[24] C. Chen, Q. Dou, H. Chen, J. Qin, and P.-A. Heng, "Synergistic image and feature adaptation: Towards cross-modality domain adaptation for medical image segmentation," in *AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 865–872.

[25] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial Intelligence and Statistics*. PMLR, 2017, pp. 1273–1282.

[26] E. Hosseini-Asl, R. Keynton, and A. El-Baz, "Alzheimer's disease diagnostics by adaptation of 3d convolutional network," in *IEEE International Conference on Image Processing*. IEEE, 2016, pp. 126–130.

[27] H. Yang, C. Chen, M. Jiang, Q. Liu, J. Cao, P. A. Heng, and Q. Dou, "Dltta: Dynamic learning rate for test-time adaptation on cross-domain medical images," *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3575–3586, 2022.

[28] J. Li, T. Chen, and X. Qian, "Generalizable pancreas segmentation modeling in ct imaging via meta-learning and latent-space feature flow generation," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 1, pp. 374–385, 2022.

[29] Z. Zhou, L. Qi, X. Yang, D. Ni, and Y. Shi, "Generalizable cross-modality medical image segmentation via style augmentation and dual normalization," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2022, pp. 20 856–20 865.

[30] L. Li, W. Ding, L. Huang, X. Zhuang, and V. Grau, "Multi-modality cardiac image computing: A survey," *Medical Image Analysis*, vol. 88, p. 102869, 2023.

[31] P. Kora, C. P. Ooi, O. Faust, U. Raghavendra, A. Gudigar, W. Y. Chan, K. Meenakshi, K. Swaraja, P. Plawiak, and U. R. Acharya, "Transfer learning techniques for medical image analysis: A review," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 1, pp. 79–107, 2022.

[32] X. Yu, J. Wang, Q.-Q. Hong, R. Teku, S.-H. Wang, and Y.-D. Zhang, "Transfer learning for medical images analyses: A survey," *Neurocomputing*, vol. 489, pp. 230–254, 2022.

[33] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: a literature review," *BMC Medical Imaging*, vol. 22, no. 1, p. 69, 2022.

[34] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[35] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2097–2106.

[36] E. Namati, J. Thiesse, J. C. Sieren, A. Ross, E. A. Hoffman, and G. McLennan, "Longitudinal assessment of lung cancer progression in the mouse using in vivo micro-ct imaging," *Medical Physics*, vol. 37, no. 9, pp. 4793–4805, 2010.

[37] J. Cai, Z. Zhang, L. Cui, Y. Zheng, and L. Yang, "Towards cross-modal organ translation and segmentation: A cycle-and shape-consistent generative adversarial network," *Medical Image Analysis*, vol. 52, pp. 174–184, 2019.

[38] K. Li, L. Yu, S. Wang, and P.-A. Heng, "Towards cross-modality medical image segmentation with online mutual knowledge distillation," in *AAAI Conference on Artificial Intelligence*, vol. 34, no. 01, 2020, pp. 775–783.

[39] L. Zhu, K. Yang, M. Zhang, L. L. Chan, T. K. Ng, and B. C. Ooi, "Semi-supervised unpaired multi-modal learning for label-efficient medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 394–404.

[40] D. Tomar, B. Bozorgtabar, M. Lortkipanidze, G. Vray, M. S. Rad, and J.-P. Thiran, "Self-supervised generative style transfer for one-

shot medical image segmentation," in *IEEE Winter Conference on Applications of Computer Vision*, 2022, pp. 1998–2008.
[41] R. Bermúdez-Chacón, P. Márquez-Neila, M. Salzmann, and P. Fua, "A domain-adaptive two-stream u-net for electron microscopy image segmentation," in *IEEE International Symposium on Biomedical Imaging*. IEEE, 2018, pp. 400–404.
[42] P. Laiz, J. Vitria, and S. Seguí, "Using the triplet loss for domain adaptation in wce," in *IEEE International Conference on Computer Vision Workshops*, 2019, pp. 0–0.
[43] N. Ghamsarian, J. Gamazo Tejero, P. Márquez-Neila, S. Wolf, M. Zinkernagel, K. Schoeffmann, and R. Sznitman, "Domain adaptation for medical image segmentation using transformation-invariant self-training," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 331–341.
[44] H. Yao, X. Hu, and X. Li, "Enhancing pseudo label quality for semi-supervised domain-generalized medical image segmentation," in *AAAI Conference on Artificial Intelligence*, vol. 36, no. 3, 2022, pp. 3099–3107.
[45] K. Li, S. Wang, L. Yu, and P.-A. Heng, "Dual-teacher: Integrating intra-domain and inter-domain teachers for annotation-efficient cardiac segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2020, pp. 418–427.
[46] K. Li, S. Wang, L. Yu, and P. A. Heng, "Dual-teacher++: Exploiting intra-domain and inter-domain knowledge with reliable transfer for cardiac segmentation," *IEEE Transactions on Medical Imaging*, vol. 40, no. 10, pp. 2771–2782, 2020.
[47] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *IEEEInternational Conference on Computer Vision*, 2017, pp. 2223–2232.
[48] Y. Zhang, S. Niu, Z. Qiu, Y. Wei, P. Zhao, J. Yao, J. Huang, Q. Wu, and M. Tan, "Covid-da: Deep domain adaptation from typical pneumonia to covid-19," *arXiv preprint arXiv:2005.01577*, 2020.
[49] X. Liu, F. Xing, N. Shusharina, R. Lim, C.-C. Jay Kuo, G. El Fakhri, and J. Woo, "Act: Semi-supervised domain-adaptive medical image segmentation with asymmetric co-training," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 66–76.
[50] H. Basak and Z. Yin, "Semi-supervised domain adaptive medical image segmentation through consistency regularized disentangled contrastive learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 260–270.
[51] H. Li, H. Liu, Y. Hu, R. Higashita, Y. Zhao, H. Qi, and J. Liu, "Restoration of cataract fundus images via unsupervised domain adaptation," in *International Symposium on Biomedical Imaging*. IEEE, 2021, pp. 516–520.
[52] C. Han, L. Rundo, K. Murao, T. Noguchi, Y. Shimahara, Z. Á. Milacski, S. Koshino, E. Sala, H. Nakayama, and S. Satoh, "Madgan: Unsupervised medical anomaly detection gan using multiple adjacent brain mri slice reconstruction," *BMC Bioinformatics*, vol. 22, no. 2, pp. 1–20, 2021.
[53] D. Tomar, M. Lortkipanidze, G. Vray, B. Bozorgtabar, and J.-P. Thiran, "Self-attentive spatial adaptive normalization for cross-modality domain adaptation," *IEEE Transactions on Medical Imaging*, vol. 40, no. 10, pp. 2926–2938, 2021.
[54] Y. Tang, Y. Tang, V. Sandfort, J. Xiao, and R. M. Summers, "Tuna-net: Task-oriented unsupervised adversarial network for disease recognition in cross-domain chest x-rays," in *Medical Image Computing and Computer Assisted Intervention*. Springer, 2019, pp. 431–440.
[55] K. Sanchez, C. Hinojosa, H. Arguello, D. Kouamé, O. Meyrignac, and A. Basarab, "Cx-dagan: Domain adaptation for pneumonia diagnosis on a small chest x-ray dataset," *IEEE Transactions on Medical Imaging*, vol. 41, no. 11, pp. 3278–3288, 2022.
[56] G. Zeng, T. D. Lerch, F. Schmaranzer, G. Zheng, J. Burger, K. Gerber, M. Tannast, K. Siebenrock, and N. Gerber, "Semantic consistent unsupervised domain adaptation for cross-modality medical image segmentation," in *Medical Image Computing and Computer Assisted Intervention*. Springer, 2021, pp. 201–210.
[57] D. Zou, Q. Zhu, and P. Yan, "Unsupervised domain adaptation with dual-scheme fusion network for medical image segmentation." in *International Joint Conference on Artificial Intelligence*, 2020, pp. 3291–3298.
[58] J. Dong, Y. Cong, G. Sun, B. Zhong, and X. Xu, "What can be transferred: Unsupervised domain adaptation for endoscopic lesions segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2020, pp. 4023–4032.
[59] C. Chen and G. Wang, "Iosuda: An unsupervised domain adaptation with input and output space alignment for joint optic disc and cup segmentation," *Applied Intelligence*, vol. 51, pp. 3880–3898, 2021.
[60] L. Peng, L. Lin, P. Cheng, Z. Huang, and X. Tang, "Unsupervised domain adaptation for cross-modality retinal vessel segmentation via disentangling representation style transfer and collaborative consistency learning," in *IEEE International Symposium on Biomedical Imaging*. IEEE, 2022, pp. 1–5.
[61] R. Wang and G. Zheng, "Cycmis: Cycle-consistent cross-domain medical image segmentation via diverse image augmentation," *Medical Image Analysis*, vol. 76, p. 102328, 2022.
[62] J. Yang, N. C. Dvornek, F. Zhang, J. Chapiro, M. Lin, and J. S. Duncan, "Unsupervised domain adaptation via disentangled representations: Application to cross-modality liver segmentation," in *Medical Image Computing and Computer Assisted Intervention*. Springer, 2019, pp. 255–263.
[63] J. Jiang and H. Veeraraghavan, "Unified cross-modality feature disentangler for unsupervised multi-domain mri abdomen organs segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2020, pp. 347–358.
[64] K. Yao, Z. Su, K. Huang, X. Yang, J. Sun, A. Hussain, and F. Coenen, "A novel 3d unsupervised domain adaptation framework for cross-modality medical image segmentation," *IEEE Journal of Biomedical and Health Informatics*, 2022.
[65] H. Zhang, J. Liu, P. Wang, Z. Yu, W. Liu, and H. Chen, "Cross-boosted multi-target domain adaptation for multi-modality histopathology image translation and segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 7, pp. 3197–3208, 2022.
[66] S. Hu, Z. Liao, and Y. Xia, "Domain specific convolution and high frequency reconstruction based unsupervised domain adaptation for medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 650–659.
[67] V. Koch, O. Holmberg, H. Spitzer, J. Schiefelbein, B. Asani, M. Hafner, and F. J. Theis, "Noise transfer for unsupervised domain adaptation of retinal oct images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 699–708.
[68] Q. Xie, Y. Li, N. He, M. Ning, K. Ma, G. Wang, Y. Lian, and Y. Zheng, "Unsupervised domain adaptation for medical image segmentation by disentanglement learning and self-training," *IEEE Transactions on Medical Imaging*, 2022.
[69] X. Du and Y. Liu, "Constraint-based unsupervised domain adaptation network for multi-modality cardiac image segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 67–78, 2021.
[70] H. Shin, H. Kim, S. Kim, Y. Jun, T. Eo, and D. Hwang, "Cosmos: Cross-modality unsupervised domain adaptation for 3d medical image segmentation based on target-aware domain translation and iterative self-training," *arXiv preprint arXiv:2203.16557*, 2022.
[71] Z. Liu, Z. Zhu, S. Zheng, Y. Liu, J. Zhou, and Y. Zhao, "Margin preserving self-paced contrastive learning towards domain adaptation for medical image segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 2, pp. 638–647, 2022.
[72] W. Feng, L. Ju, L. Wang, K. Song, X. Zhao, and Z. Ge, "Unsupervised domain adaptation for medical image segmentation by selective entropy constraints and adaptive semantic alignment," in *AAAI Conference on Artificial Intelligence*, vol. 37, no. 1, 2023, pp. 623–631.
[73] C. Li, D. Liu, H. Li, Z. Zhang, G. Lu, X. Chang, and W. Cai, "Domain adaptive nuclei instance segmentation and classification via category-aware feature alignment and pseudo-labelling," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 715–724.
[74] H. Cho, K. Nishimura, K. Watanabe, and R. Bise, "Effective pseudo-labeling based on heatmap for unsupervised domain adaptation in cell detection," *Medical Image Analysis*, vol. 79, p. 102436, 2022.
[75] X. Liu, F. Xing, M. Stone, J. Zhuo, T. Reese, J. L. Prince, G. El Fakhri, and J. Woo, "Generative self-training for cross-domain unsupervised tagged-to-cine mri synthesis," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 138–148.
[76] Q. Dou, C. Ouyang, C. Chen, H. Chen, B. Glocker, X. Zhuang, and P.-A. Heng, "Pnp-adanet: Plug-and-play adversarial domain adaptation network at unpaired cross-modality cardiac segmentation," *IEEE Access*, vol. 7, pp. 99 065–99 076, 2019.
[77] K. Jiang, L. Quan, and T. Gong, "Disentangled representation and cross-modality image translation based unsupervised domain adaptation method for abdominal organ segmentation," *International Journal*

of Computer Assisted Radiology and Surgery*, vol. 17, no. 6, pp. 1101–1113, 2022.

[78] B. Liu, D. Pan, Z. Shuai, and H. Song, “Ecsd-net: A joint optic disc and cup segmentation and glaucoma classification network based on unsupervised domain adaptation,” *Computer Methods and Programs in Biomedicine*, vol. 213, p. 106530, 2022.

[79] H. Lei, W. Liu, H. Xie, B. Zhao, G. Yue, and B. Lei, “Unsupervised domain adaptation based image synthesis and feature alignment for joint optic disc and cup segmentation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 90–102, 2021.

[80] W. Yan, Y. Wang, M. Xia, and Q. Tao, “Edge-guided output adaptor: Highly efficient adaptation module for cross-vendor medical image segmentation,” *IEEE Signal Processing Letters*, vol. 26, no. 11, pp. 1593–1597, 2019.

[81] R. Shen, J. Yao, K. Yan, K. Tian, C. Jiang, and K. Zhou, “Unsupervised domain adaptation with adversarial learning for mass detection in mammogram,” *Neurocomputing*, vol. 393, pp. 27–37, 2020.

[82] C. Chen, Q. Dou, H. Chen, J. Qin, and P. A. Heng, “Unsupervised bidirectional cross-modality adaptation via deeply synergistic image and feature alignment for medical image segmentation,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 7, pp. 2494–2505, 2020.

[83] S. Y. Shin, S. Lee, and R. M. Summers, “Unsupervised domain adaptation for small bowel segmentation using disentangled representation,” in *Medical Image Computing and Computer Assisted Intervention*. Springer, 2021, pp. 282–292.

[84] H. Liu, Y. Zhuang, E. Song, X. Xu, and C.-C. Hung, “A bidirectional multilayer contrastive adaptation network with anatomical structure preservation for unpaired cross-modality medical image segmentation,” *Computers in Biology and Medicine*, vol. 149, p. 105964, 2022.

[85] Z. Su, K. Yao, X. Yang, Q. Wang, Y. Yan, J. Sun, and K. Huang, “Mind the gap: Alleviating local imbalance for unsupervised cross-modality medical image segmentation,” *IEEE Journal of Biomedical and Health Informatics*, 2023.

[86] M. G. B. Calisto and S. K. Lai-Yuen, “C-mada: unsupervised cross-modality adversarial domain adaptation framework for medical image segmentation,” in *Medical Imaging 2022: Image Processing*, vol. 12032. SPIE, 2022, pp. 971–978.

[87] X. Bian, X. Luo, C. Wang, W. Liu, and X. Lin, “Dda-net: Unsupervised cross-modality medical image segmentation via dual domain adaptation,” *Computer Methods and Programs in Biomedicine*, vol. 213, p. 106531, 2022.

[88] D. D. Pham, S. Koesnadi, G. Dovletov, and J. Pauli, “Unsupervised adversarial domain adaptation for multi-label classification of chest x-ray,” in *IEEE International Symposium on Biomedical Imaging*. IEEE, 2021, pp. 1236–1240.

[89] C. Yoo, H. W. Lee, and J.-W. Kang, “Transferring structured knowledge in unsupervised domain adaptation of a sleep staging network,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 3, pp. 1273–1284, 2021.

[90] J. Liu, H. Liu, S. Gong, Z. Tang, Y. Xie, H. Yin, and J. P. Niyoyita, “Automated cardiac segmentation of cross-modal medical images using unsupervised multi-domain adaptation and spatial neural attention structure,” *Medical Image Analysis*, vol. 72, p. 102135, 2021.

[91] X. Chen, T. Kuang, H. Deng, S. H. Fung, J. Gateno, J. J. Xia, and P.-T. Yap, “Dual adversarial attention mechanism for unsupervised domain adaptive medical image segmentation,” *IEEE Transactions on Medical Imaging*, vol. 41, no. 11, pp. 3445–3453, 2022.

[92] C. S. Perone, P. Ballester, R. C. Barros, and J. Cohen-Adad, “Unsupervised domain adaptation for medical imaging segmentation with self-ensembling,” *NeuroImage*, vol. 194, pp. 1–11, 2019.

[93] S. Kuang, H. C. Woodruff, R. Granzier, T. J. van Nijnatten, M. B. Lobbes, M. L. Smidt, P. Lambin, and S. Mehrkanoon, “Mscda: Multi-level semantic-guided contrast improves unsupervised domain adaptation for breast mri segmentation in small datasets,” *Neural Networks*, vol. 165, pp. 119–134, 2023.

[94] V. Srivastav, A. Gangi, and N. Padoy, “Unsupervised domain adaptation for clinician pose estimation and instance segmentation in the operating room,” *Medical Image Analysis*, vol. 80, p. 102525, 2022.

[95] Z. Shanis, S. Gerber, M. Gao, and A. Enquobahrie, “Intramodality domain adaptation using self ensembling and adversarial training,” in *MICCAI Workshop on Domain Adaptation and Representation Transfer*. Springer, 2019, pp. 28–36.

[96] P. Liu, B. Kong, Z. Li, S. Zhang, and R. Fang, “Cfea: Collaborative feature ensembling adaptation for domain adaptation in unsupervised optic disc and cup segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 521–529.

[97] Z. Zhao, F. Zhou, K. Xu, Z. Zeng, C. Guan, and S. K. Zhou, “Le-uda: Label-efficient unsupervised domain adaptation for medical image segmentation,” *IEEE Transactions on Medical Imaging*, vol. 42, no. 3, pp. 633–646, 2022.

[98] Z. Zhao, K. Xu, S. Li, Z. Zeng, and C. Guan, “Mt-uda: Towards unsupervised cross-modality medical image segmentation with limited source labels,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 293–303.

[99] S. Liu, S. Yin, L. Qu, and M. Wang, “Reducing domain gap in frequency and spatial domain for cross-modality domain adaptation on medical image segmentation,” in *AAAI Conference on Artificial Intelligence*, vol. 37, no. 2, 2023, pp. 1719–1727.

[100] J. Hong, S. C.-H. Yu, and W. Chen, “Unsupervised domain adaptation for cross-modality liver segmentation via joint adversarial learning and self-learning,” *Applied Soft Computing*, vol. 121, p. 108729, 2022.

[101] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016.

[102] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, “Graph attention networks,” *arXiv preprint arXiv:1710.10903*, 2017.

[103] J. Guo, K. Huang, R. Zhang, and X. Yi, “Es-gnn: Generalizing graph neural networks beyond homophily with edge splitting,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.

[104] J. Guo, K. Huang, X. Yi, and R. Zhang, “Learning disentangled graph convolutional networks locally and globally,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 3, pp. 3640–3651, 2022.

[105] D. Mahapatra, S. Korevaar, B. Bozorgtabar, and R. Tennakoon, “Unsupervised domain adaptation using feature disentanglement and gcns for medical image classification,” in *European Conference on Computer Vision*. Springer, 2022, pp. 735–748.

[106] K. Gopinath, C. Desrosiers, and H. Lombaert, “Graph domain adaptation for alignment-invariant brain surface segmentation,” in *MICCAI Workshop on Uncertainty for Safe Utilization of Machine Learning in Medical Imaging*. Springer, 2020, pp. 152–163.

[107] Y. Fang, M. Wang, G. G. Potter, and M. Liu, “Unsupervised cross-domain functional mri adaptation for automated major depressive disorder identification,” *Medical Image Analysis*, vol. 84, p. 102707, 2023.

[108] Y. Zhang, N. Yuan, Z. Zhang, J. Du, T. Wang, B. Liu, A. Yang, K. Lv, G. Ma, and B. Lei, “Unsupervised domain selective graph convolutional network for preoperative prediction of lymph node metastasis in gastric cancer,” *Medical Image Analysis*, vol. 79, p. 102467, 2022.

[109] X. Liu, X. Liu, B. Hu, W. Ji, F. Xing, J. Lu, J. You, C.-C. J. Kuo, G. El Fakhri, and J. Woo, “Subtype-aware unsupervised domain adaptation for medical diagnosis,” in *AAAI Conference on Artificial Intelligence*, vol. 35, no. 3, 2021, pp. 2189–2197.

[110] J. Li, C. Feng, X. Lin, and X. Qian, “Utilizing gcn and meta-learning strategy in unsupervised domain adaptation for pancreatic cancer segmentation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 79–89, 2021.

[111] J. Liu, X. Guo, and Y. Yuan, “Prototypical interaction graph for unsupervised domain adaptation in surgical instrument segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 272–281.

[112] ——, “Graph-based surgical instrument adaptive segmentation via domain-common knowledge,” *IEEE Transactions on Medical Imaging*, vol. 41, no. 3, pp. 715–726, 2021.

[113] M. Long, H. Zhu, J. Wang, and M. I. Jordan, “Unsupervised domain adaptation with residual transfer networks,” *arXiv preprint arXiv:1602.04433*, 2016.

[114] J. Gao, Q. Lao, Q. Kang, P. Liu, L. Zhang, and K. Li, “Unsupervised cross-disease domain adaptation by lesion scale matching,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 660–670.

[115] Q. Hu, H. Li, and J. Zhang, “Domain-adaptive 3d medical image synthesis: An efficient unsupervised approach,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 495–504.

[116] C. Lu, S. Zheng, and G. Gupta, “Unsupervised domain adaptation for cardiac segmentation: Towards structure mutual information maximization,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2022, pp. 2588–2597.

[117] X. Liu, S. Li, Y. Ge, P. Ye, J. You, and J. Lu, “Ordinal unsupervised domain adaptation with recursively conditional gaussian imposed variational disentanglement,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[118] B. Xu, K. Wu, Y. Wu, J. He, and C. Chen, "Dynamic adversarial domain adaptation based on multikernel maximum mean discrepancy for breast ultrasound image classification," *Expert Systems with Applications*, vol. 207, p. 117978, 2022.

[119] F. Wu and X. Zhuang, "Cf distance: a new domain discrepancy metric and application to explicit domain adaptation for cross-modality cardiac image segmentation," *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 4274–4285, 2020.

[120] P. Sager, S. Salzmann, F. Burn, and T. Stadelmann, "Unsupervised domain adaptation for vertebrae detection and identification in 3d ct volumes using a domain sanity loss," *Journal of Imaging*, vol. 8, no. 8, 2022.

[121] E. S. Lubana, C. I. Tang, F. Kawsar, R. P. Dick, and A. Mathur, "Orchestra: Unsupervised federated learning via globally consistent clustering," *arXiv preprint arXiv:2205.11506*, 2022.

[122] H. Kim, Y. Kwak, M. Jung, J. Shin, Y. Kim, and C. Kim, "Protofl: Unsupervised federated learning via prototypical distillation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 6470–6479.

[123] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas, "Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation," in *International MICCAI Brainlesion Workshop*. Springer, 2019, pp. 92–104.

[124] Q. Dou, T. Y. So, M. Jiang, Q. Liu, V. Vardhanabhuti, G. Kaissis, Z. Li, W. Si, H. H. Lee, K. Yu *et al.*, "Federated deep learning for detecting covid-19 lung abnormalities in ct: a privacy-preserving multinational validation study," *NPJ Digital Medicine*, vol. 4, no. 1, p. 60, 2021.

[125] I. Feki, S. Ammar, Y. Kessentini, and K. Muhammad, "Federated learning for covid-19 screening from chest x-ray images," *Applied Soft Computing*, vol. 106, p. 107330, 2021.

[126] G. Elmas, S. U. Dar, Y. Korkmaz, E. Ceyani, B. Susam, M. Ozbey, S. Avestimehr, and T. Çukur, "Federated learning of generative image priors for mri reconstruction," *IEEE Transactions on Medical Imaging*, 2022.

[127] J. Lo, T. Y. Timothy, D. Ma, P. Zang, J. P. Owen, Q. Zhang, R. K. Wang, M. F. Beg, A. Y. Lee, Y. Jia *et al.*, "Federated learning for microvasculature segmentation and diabetic retinopathy classification of oct data," *Ophthalmology Science*, vol. 1, no. 4, p. 100069, 2021.

[128] L. Li, N. Xie, and S. Yuan, "A federated learning framework for breast cancer histopathological image classification," *Electronics*, vol. 11, no. 22, p. 3767, 2022.

[129] M. Adnan, S. Kalra, J. C. Cresswell, G. W. Taylor, and H. R. Tizhoosh, "Federated learning and differential privacy for medical image analysis," *Scientific Reports*, vol. 12, no. 1, p. 1953, 2022.

[130] B. C. Tedeschini, S. Savazzi, R. Stoklasa, L. Barbieri, I. Stathopoulos, M. Nicoli, and L. Serio, "Decentralized federated learning for healthcare networks: A case study on tumor segmentation," *IEEE Access*, vol. 10, pp. 8693–8708, 2022.

[131] S. Yang, H. Hwang, D. Kim, R. Dua, J.-Y. Kim, E. Yang, and E. Choi, "Towards the practical utility of federated learning in the medical domain," *arXiv preprint arXiv:2207.03075*, 2022.

[132] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," *Proceedings of Machine learning and systems*, vol. 2, pp. 429–450, 2020.

[133] L. Collins, H. Hassani, A. Mokhtari, and S. Shakkottai, "Exploiting shared representations for personalized federated learning," in *International Conference on Machine Learning*. PMLR, 2021, pp. 2089–2099.

[134] X. Liao, J. Zhou, and J. Shu, "A blockchain enabled federal domain generalization based architecture for dependable medical image segmentation," in *IEEE Advanced Information Technology, Electronic and Automation Control Conference*. IEEE, 2022, pp. 1655–1658.

[135] M. Jiang, Z. Wang, and Q. Dou, "Harmofl: Harmonizing local and global drifts in federated learning on heterogeneous medical images," in *AAAI Conference on Artificial Intelligence*, vol. 36, no. 1, 2022, pp. 1087–1095.

[136] A. Linardos, K. Kushibar, S. Walsh, P. Gkontra, and K. Lekadir, "Federated learning for multi-center imaging diagnostics: a simulation study in cardiovascular disease," *Scientific Reports*, vol. 12, no. 1, p. 3551, 2022.

[137] A. E. Cetinkaya, M. Akin, and S. Sagiroglu, "Improving performance of federated learning based medical image analysis in non-iid settings using image augmentation," in *International Conference on Information Security and Cryptology*. IEEE, 2021, pp. 69–74.

[138] Z. Yan, J. Wicaksana, Z. Wang, X. Yang, and K.-T. Cheng, "Variation-aware federated learning with multi-source decentralized medical image data," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 7, pp. 2615–2628, 2020.

[139] M. Manthe, S. Duffner, and C. Lartizien, "Whole brain radiomics for clustered federated personalization in brain tumor segmentation," in *Medical Imaging with Deep Learning*. PMLR, 2024, pp. 957–977.

[140] L. Yi, J. Zhang, R. Zhang, J. Shi, G. Wang, and X. Liu, "Su-net: an efficient encoder-decoder model of federated learning for brain tumor segmentation," in *International Conference on Artificial Neural Networks*. Springer, 2020, pp. 761–773.

[141] B. Lei, Y. Zhu, E. Liang, P. Yang, S. Chen, H. Hu, H. Xie, Z. Wei, F. Hao, X. Song *et al.*, "Federated domain adaptation via transformer for multi-site alzheimer's disease diagnosis," *IEEE Transactions on Medical Imaging*, 2023.

[142] P. Guo, P. Wang, J. Zhou, S. Jiang, and V. M. Patel, "Multi-institutional collaborations for improving deep learning-based magnetic resonance image reconstruction using federated learning," in *IEEE Conference on Computer Vision and Pttern Recognition*, 2021, pp. 2423–2432.

[143] O. Dalmaz, U. Mirza, G. Elmas, M. Özbey, S. U. Dar, and T. Çukur, "A specificity-preserving generative model for federated mri translation," in *International Workshop on Distributed, Collaborative, and Federated Learning*. Springer, 2022, pp. 79–88.

[144] N. Wu, L. Yu, X. Yang, K.-T. Cheng, and Z. Yan, "Fediic: Towards robust federated learning for class-imbalanced medical image classification," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 692–702.

[145] Z. Liu, F. Wu, Y. Wang, M. Yang, and X. Pan, "Fedcl: Federated contrastive learning for multi-center medical image classification," *Pattern Recognition*, vol. 143, p. 109739, 2023.

[146] J. Wang, Y. Jin, D. Stoyanov, and L. Wang, "Feddp: Dual personalization in federated medical image segmentation," *IEEE Transactions on Medical Imaging*, 2023.

[147] J. Wang, Y. Jin, and L. Wang, "Personalizing federated medical image segmentation via local calibration," in *European Conference on Computer Vision*. Springer, 2022, pp. 456–472.

[148] A. Xu, W. Li, P. Guo, D. Yang, H. R. Roth, A. Hatamizadeh, C. Zhao, D. Xu, H. Huang, and Z. Xu, "Closing the generalization gap of cross-silo federated medical image segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2022, pp. 20 866–20 875.

[149] X. Xu, H. H. Deng, T. Chen, T. Kuang, J. C. Barber, D. Kim, J. Gateno, J. J. Xia, and P. Yan, "Federated cross learning for medical image segmentation," in *Medical Imaging with Deep Learning*. PMLR, 2024, pp. 1441–1452.

[150] G. N. Gunesli, M. Bilal, S. E. A. Raza, and N. M. Rajpoot, "A federated learning approach to tumor detection in colon histology images," *Journal of Medical Systems*, vol. 47, no. 1, p. 99, 2023.

[151] W. Ding, M. Abdel-Basset, H. Hawash, M. Pratama, and W. Pedrycz, "Generalizable segmentation of covid-19 infection from multi-site tomography scans: a federated learning framework," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2023.

[152] X. Li, Y. Gu, N. Dvornek, L. H. Staib, P. Ventola, and J. S. Duncan, "Multi-site fmri analysis using privacy-preserving federated learning and domain adaptation: Abide results," *Medical Image Analysis*, vol. 65, p. 101765, 2020.

[153] M. Zhang, L. Qu, P. Singh, J. Kalpathy-Cramer, and D. L. Rubin, "Splitavg: A heterogeneity-aware federated deep learning method for medical imaging," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 9, pp. 4635–4644, 2022.

[154] S. Otálora, J. Rafael-Patiño, A. Madrona, E. Fischi-Gomez, V. Ravano, T. Kober, S. Christensen, A. Hakim, R. Wiest, J. Richiardi *et al.*, "Weighting schemes for federated learning in heterogeneous and imbalanced segmentation datasets," in *International MICCAI Brainlesion Workshop*. Springer, 2022, pp. 45–56.

[155] C. Shen, P. Wang, H. R. Roth, D. Yang, D. Xu, M. Oda, W. Wang, C.-S. Fuh, P.-T. Chen, K.-L. Liu *et al.*, "Multi-task federated learning for heterogeneous pancreas segmentation," in *Workshop on Clinical Image-Based Procedures, Distributed and Collaborative Learning, Artificial Intelligence for Combating COVID-19 and Secure and Privacy-Preserving Machine Learning*. Springer, 2021, pp. 101–110.

[156] Z. Deng, D. Li, S. Tan, Y. Fu, X. Yuan, X. Huang, Y. Zhang, and G. Zhou, "Fedgrav: An adaptive federated aggregation algorithm for multi-institutional medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 170–180.

[157] M. I. Khan, M. A. Azeem, E. Alhoniemi, E. Kontio, S. A. Khan, and M. Jafaritadi, "Regularized weight aggregation in networked federated learning for glioblastoma segmentation," in *International MICCAI Brainlesion Workshop*. Springer, 2022, pp. 121–132.

[158] Y. Yeganeh, A. Farshad, N. Navab, and S. Albarqouni, "Inverse distance aggregation for federated learning with non-iid data," in *MICCAI Workshop on Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*. Springer, 2020, pp. 150–159.

[159] R. Zhang, Z. Fan, Q. Xu, J. Yao, Y. Zhang, and Y. Wang, "Grace: A generalized and personalized federated learning method for medical imaging," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 14–24.

[160] G. Yue, P. Wei, T. Zhou, Y. Song, C. Zhao, T. Wang, and B. Lei, "Specificity-aware federated learning with dynamic feature fusion network for imbalanced medical image classification," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[161] Y. Yin, H. Yang, Q. Liu, M. Jiang, C. Chen, Q. Dou, and P.-A. Heng, "Efficient federated tumor segmentation via normalized tensor aggregation and client pruning," in *International MICCAI Brainlesion Workshop*. Springer, 2021, pp. 433–443.

[162] M. Jiang, H. Yang, C. Cheng, and Q. Dou, "Iop-fl: inside-outside personalization for federated medical image segmentation," *IEEE Transactions on Medical Imaging*, 2023.

[163] Z. Chen, M. Zhu, C. Yang, and Y. Yuan, "Personalized retrogress-resilient framework for real-world medical federated learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 347–356.

[164] S. M. Hosseini, M. Sikaroudi, M. Babaie, and H. Tizhoosh, "Proportionally fair hospital collaborations in federated learning of histopathology images," *IEEE Transactions on Medical Imaging*, 2023.

[165] H. Wu, B. Zhang, C. Chen, and J. Qin, "Federated semi-supervised medical image segmentation via prototype-based pseudo-labeling and contrastive learning," *IEEE Transactions on Medical Imaging*, 2023.

[166] L. Qiu, J. Cheng, H. Gao, W. Xiong, and H. Ren, "Federated semi-supervised learning for medical image segmentation via pseudo-label denoising," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[167] W. Liu, J. Mo, and F. Zhong, "Class imbalanced medical image classification based on semi-supervised federated learning," *Applied Sciences*, vol. 13, no. 4, p. 2109, 2023.

[168] Q. Liu, H. Yang, Q. Dou, and P.-A. Heng, "Federated semi-supervised medical image classification via inter-client relation matching," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 325–335.

[169] R. Yan, L. Qu, Q. Wei, S.-C. Huang, L. Shen, D. Rubin, L. Xing, and Y. Zhou, "Label-efficient self-supervised federated learning for tackling data heterogeneity in medical imaging," *IEEE Transactions on Medical Imaging*, 2023.

[170] M. Zhu, Z. Chen, and Y. Yuan, "Feddm: Federated weakly supervised segmentation via annotation calibration and gradient de-conflicting," *IEEE Transactions on Medical Imaging*, 2023.

[171] H. Kassem, D. Alapatt, P. Mascagni, C. AI4SafeChole, A. Karargyris, and N. Padoy, "Federated cycling (fedcy): Semi-supervised federated learning of surgical phases," *IEEE Transactions on Medical Imaging*, 2022.

[172] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[173] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, 2012.

[174] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition*. Ieee, 2009, pp. 248–255.

[175] N. M. Khan, N. Abraham, and M. Hon, "Transfer learning with intelligent training data selection for prediction of alzheimer's disease," *IEEE Access*, vol. 7, pp. 72 726–72 735, 2019.

[176] V. Chouhan, S. K. Singh, A. Khamparia, D. Gupta, P. Tiwari, C. Moreira, R. Damaševičius, and V. H. C. De Albuquerque, "A novel transfer learning based approach for pneumonia detection in chest x-ray images," *Applied Sciences*, vol. 10, no. 2, p. 559, 2020.

[177] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Detrac: Transfer learning of class decomposed medical images in convolutional neural networks," *IEEE Access*, vol. 8, pp. 74 901–74 913, 2020.

[178] Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, and J. Lu, "Brain tumor classification for mr images using transfer learning and fine-tuning," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 34–46, 2019.

[179] S. Chen, K. Ma, and Y. Zheng, "Med3d: Transfer learning for 3d medical image analysis," *arXiv preprint arXiv:1904.00625*, 2019.

[180] R. Venkataramani, H. Ravishankar, and S. Anamandra, "Towards continuous domain adaptation for medical imaging," in *IEEE International Symposium on Biomedical Imaging*. IEEE, 2019, pp. 443–446.

[181] M. Lenga, H. Schulz, and A. Saalbach, "Continual learning for domain adaptation in chest x-ray classification," in *Medical Imaging with Deep Learning*. PMLR, 2020, pp. 413–423.

[182] N. Karani, K. Chaitanya, C. Baumgartner, and E. Konukoglu, "A lifelong learning approach to brain mr segmentation across scanners and protocols," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 476–484.

[183] B. Chidlovskii, S. Clinchant, and G. Csurka, "Domain adaptation in the absence of source domain data," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 451–460.

[184] D. Wang, E. Shelhamer, S. Liu, B. Olshausen, and T. Darrell, "Tent: Fully test-time adaptation by entropy minimization," in *International Conference on Learning Representations*, 2021.

[185] J. Wu, G. Wang, R. Gu, T. Lu, Y. Chen, W. Zhu, T. Vercauteren, S. Ourselin, and S. Zhang, "Upl-sfda: Uncertainty-aware pseudo label guided source-free domain adaptation for medical image segmentation," *IEEE Transactions on Medical Imaging*, 2023.

[186] K.-S. Cheng, Q.-W. Zhang, H.-W. Tsai, N.-t. Li, and P.-C. Chung, "Domain-centroid-guided progressive teacher-based knowledge distillation for source-free domain adaptation of histopathological images," *IEEE Transactions on Artificial Intelligence*, 2023.

[187] W. Zhou, J. Ji, W. Cui, and Y. Yi, "Pseudo-label clustering-driven dual-level contrast learning based source-free domain adaptation for fundus image segmentation," in *Chinese Conference on Pattern Recognition and Computer Vision*. Springer, 2023, pp. 492–503.

[188] Q. Yu, N. Xi, J. Yuan, Z. Zhou, K. Dang, and X. Ding, "Source-free domain adaptation for medical image segmentation via prototype-anchored feature alignment and contrastive learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 3–12.

[189] C. Li, W. Chen, X. Luo, Y. He, and Y. Tan, "Adaptive pseudo labeling for source-free domain adaptation in medical image segmentation," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2022, pp. 1091–1095.

[190] C. Chen, Q. Liu, Y. Jin, Q. Dou, and P.-A. Heng, "Source-free domain adaptive fundus image segmentation with denoised pseudo-labeling," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 225–235.

[191] C. Yang, X. Guo, Z. Chen, and Y. Yuan, "Source free domain adaptation for medical image segmentation with fourier style mining," *Medical Image Analysis*, vol. 79, p. 102457, 2022.

[192] H. Li, Z. Lin, Z. Qiu, Z. Li, K. Niu, N. Guo, H. Fu, Y. Hu, and J. Liu, "Enhancing and adapting in the clinic: Source-free unsupervised domain adaptation for medical image enhancement," *IEEE Transactions on Medical Imaging*, 2023.

[193] X. Huang, X. Yang, H. Dou, Y. Huang, L. Zhang, Z. Liu, Z. Yan, L. Liu, Y. Zou, X. Hu *et al.*, "Test-time bi-directional adaptation between image and model for robust segmentation," *Computer Methods and Programs in Biomedicine*, vol. 233, p. 107477, 2023.

[194] S. Qiu, "Causality-inspired source-free domain adaptation for medical image classification," in *International Conference on Image and Graphics*. Springer, 2023, pp. 68–80.

[195] Y. Ye, Z. Liu, Y. Zhang, J. Li, and H. Shen, "Alleviating style sensitivity then adapting: Source-free domain adaptation for medical image segmentation," in *ACM International Conference on Multimedia*, 2022, pp. 1935–1944.

[196] D. Tomar, G. M. G. Vray, B. Bozorgtabar, and J.-P. Thiran, "Opttta: Learnable test-time augmentation for source-free medical image segmentation under domain shift," in *International Conference on Medical Imaging with Deep Learning*. PMLR, 2022, pp. 1192–1217.

[197] Y. He, A. Carass, L. Zuo, B. E. Dewey, and J. L. Prince, "Autoencoder based self-supervised test-time adaptation for medical image analysis," *Medical Image Analysis*, vol. 72, p. 102136, 2021.

[198] C. Zhou, W. Zhang, H. Chen, and L. Chen, "Domain adaptation for medical image classification without source data," in *IEEE International Conference on Bioinformatics and Biomedicine*. IEEE, 2022, pp. 2224–2230.

[199] S. Hu, Z. Liao, and Y. Xia, "Prosfda: Prompt learning based source-free domain adaptation for medical image segmentation," *arXiv preprint arXiv:2211.11514*, 2022.

[200] X. Liu, F. Xing, C. Yang, G. El Fakhri, and J. Woo, "Adapting off-the-shelf source segmenter for target medical image segmentation," in

International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2021, pp. 549–559.

[201] X. Liu, F. Xing, G. El Fakhri, and J. Woo, "Memory consistent unsupervised off-the-shelf model adaptation for source-relaxed medical image segmentation," *Medical Image Analysis*, vol. 83, p. 102641, 2023.

[202] Z. Su, J. Guo, K. Yao, X. Yang, Q. Wang, and K. Huang, "Unraveling batch normalization for realistic test-time adaptation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 13, 2024, pp. 15 136–15 144.

[203] Y. Feng, Y. Luo, and J. Yang, "Cross-platform privacy-preserving ct image covid-19 diagnosis based on source-free domain adaptation," *Knowledge-Based Systems*, vol. 264, p. 110324, 2023.

[204] W. Ma, C. Chen, S. Zheng, J. Qin, H. Zhang, and Q. Dou, "Test-time adaptation with calibration of medical image classification nets for label distribution shift," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 313–323.

[205] J. Wu, R. Gu, T. Lu, S. Zhang, and G. Wang, "Upl-tta: Uncertainty-aware pseudo label guided fully test time adaptation for fetal brain segmentation," in *International Conference on Information Processing in Medical Imaging*. Springer, 2023, pp. 237–249.

[206] R. Wen, H. Yuan, D. Ni, W. Xiao, and Y. Wu, "From denoising training to test-time adaptation: Enhancing domain generalization for medical image segmentation," in *IEEE Winter Conference on Applications of Computer Vision*, 2024, pp. 464–474.

[207] N. Karani, E. Erdil, K. Chaitanya, and E. Konukoglu, "Test-time adaptable neural networks for robust medical image segmentation," *Medical Image Analysis*, vol. 68, p. 101907, 2021.

[208] Y. Zhang, T. Zhou, Y. Tao, S. Wang, Y. Wu, B. Liu, P. Gu, Q. Chen, and D. Z. Chen, "Testfit: A plug-and-play one-pass test time method for medical image segmentation," *Medical Image Analysis*, vol. 92, p. 103069, 2024.

[209] J. Hong, Y.-D. Zhang, and W. Chen, "Source-free unsupervised domain adaptation for cross-modality abdominal multi-organ segmentation," *Knowledge-Based Systems*, vol. 250, p. 109155, 2022.

[210] M. Bateson, H. Kervadec, J. Dolz, H. Lombaert, and I. B. Ayed, "Source-free domain adaptation for image segmentation," *Medical Image Analysis*, vol. 82, p. 102617, 2022.

[211] M. Bateson, H. Lombaert, and I. Ben Ayed, "Test-time adaptation with shape moments for image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 736–745.

[212] Y. Zhang, K. Huang, C. Chen, Q. Chen, and P.-A. Heng, "Satta: Semantic-aware test-time adaptation for cross-domain medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 148–158.

[213] Q. Liu, C. Chen, Q. Dou, and P.-A. Heng, "Single-domain generalization in medical image segmentation via test-time adaptation from shape dictionary," in *AAAI Conference on Artificial Intelligence*, vol. 36, no. 2, 2022, pp. 1756–1764.

[214] J. Zhu, B. Bolsterlee, B. V. Chow, Y. Song, and E. Meijering, "Uncertainty and shape-aware continual test-time adaptation for cross-domain segmentation of medical images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 659–669.

[215] A. Bigalke, L. Hansen, J. Diesel, C. Hennigs, P. Rostalski, and M. P. Heinrich, "Anatomy-guided domain adaptation for 3d in-bed human pose estimation," *Medical Image Analysis*, vol. 89, p. 102887, 2023.

[216] M. Hu, T. Song, Y. Gu, X. Luo, J. Chen, Y. Chen, Y. Zhang, and S. Zhang, "Fully test-time adaptation for image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 251–260.

[217] Z. Li, J. Yang, Y. Xu, L. Zhang, W. Dong, and B. Du, "Scale-aware test-time click adaptation for pulmonary nodule and mass segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 681–691.

[218] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-learning in neural networks: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5149–5169, 2021.

[219] Q. Liu, C. Chen, J. Qin, Q. Dou, and P.-A. Heng, "Feddg: Federated domain generalization on medical image segmentation via episodic learning in continuous frequency space," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2021, pp. 1013–1023.

[220] D. D. Pham, G. Dovletov, and J. Pauli, "Liver segmentation in ct with mri data: zero-shot domain adaptation by contour extraction and shape priors," in *IEEE International Symposium on Biomedical Imaging*. IEEE, 2020, pp. 1538–1542.

[221] H. Li, Y. Wang, R. Wan, S. Wang, T.-Q. Li, and A. Kot, "Domain generalization for medical imaging classification with linear-dependency regularization," *Advances in Neural Information Processing Systems*, vol. 33, pp. 3118–3129, 2020.

[222] Q. Dou, D. Coelho de Castro, K. Kamnitsas, and B. Glocker, "Domain generalization via model-agnostic learning of semantic features," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[223] L. Zhang, X. Wang, D. Yang, T. Sanford, S. Harmon, B. Turkbey, B. J. Wood, H. Roth, A. Myronenko, D. Xu *et al.*, "Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation," *IEEE Transactions on Medical Imaging*, vol. 39, no. 7, pp. 2531–2540, 2020.

[224] R. A. Kamraoui, V.-T. Ta, T. Tourdias, B. Mansencal, J. V. Manjon, and P. Coupé, "Deeplesionbrain: Towards a broader deep-learning generalization for multiple sclerosis lesion segmentation," *Medical Image Analysis*, vol. 76, p. 102312, 2022.

[225] C. Ouyang, C. Chen, S. Li, Z. Li, C. Qin, W. Bai, and D. Rueckert, "Causality-inspired single-source domain generalization for medical image segmentation," *IEEE Transactions on Medical Imaging*, vol. 42, no. 4, pp. 1095–1106, 2022.

[226] Z. Su, K. Yao, X. Yang, K. Huang, Q. Wang, and J. Sun, "Rethinking data augmentation for single-source domain generalization in medical image segmentation," in *AAAI Conference on Artificial Intelligence*, vol. 37, no. 2, 2023, pp. 2366–2374.

[227] S. Hu, Z. Liao, and Y. Xia, "Devil is in channels: Contrastive single domain generalization for medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 14–23.

[228] C. Chen, Z. Li, C. Ouyang, M. Sinclair, W. Bai, and D. Rueckert, "Maxstyle: Adversarial style composition for robust medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 151–161.

[229] Y. Xu, S. Xie, M. Reynolds, M. Ragoza, M. Gong, and K. Batmanghelich, "Adversarial consistency for single domain generalization in medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 671–681.

[230] C. Chen, K. Hammernik, C. Ouyang, C. Qin, W. Bai, and D. Rueckert, "Cooperative training and latent space data augmentation for robust medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 149–159.

[231] Y. Zhang, X. Li, H. Chen, A. L. Yuille, Y. Liu, and Z. Zhou, "Continual learning for abdominal multi-organ and tumor segmentation," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2023, pp. 35–45.

[232] C. González, A. Ranem, D. Pinto dos Santos, A. Othman, and A. Mukhopadhyay, "Lifelong nnu-net: a framework for standardized medical continual learning," *Scientific Reports*, vol. 13, no. 1, p. 9381, 2023.

[233] Z. Wang, Z. Wu, D. Agarwal, and J. Sun, "Medclip: Contrastive learning from unpaired medical images and text," *arXiv preprint arXiv:2210.10163*, 2022.

[234] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo *et al.*, "Segment anything," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.

[235] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*. PMLR, 2021, pp. 8748–8763.

[236] S. Aleem, F. Wang, M. Maniparambil, E. Arazo, J. Dietlmeier, K. Curran, N. E. Connor, and S. Little, "Test-time adaptation with salip: A cascade of sam and clip for zero-shot medical image segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 5184–5193.

[237] Y. Zhao, X. Wang, T. Che, G. Bao, and S. Li, "Multi-task deep learning for medical image computing and analysis: A review," *Computers in Biology and Medicine*, vol. 153, p. 106496, 2023.

[238] Y. Cao, L. Cui, L. Zhang, F. Yu, Z. Li, and Y. Xu, "Mmtn: multi-modal memory transformer network for image-report consistent medical report generation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 1, 2023, pp. 277–285.