MLN-net: A multi-source medical image segmentation method for clustered microcalcifications using multiple layer normalization

Ke Wang[a,b], Zanting Ye[c,*], Xiang Xie[d], Haidong Cui[e], Tao Chen[f] and Banteng Liu[a,c,*]

[a]College of Information Science and Technology, Zhejiang Shuren University, Hangzhou, China
[b]State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China
[c]School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, China
[d]School of Engineering, Newcastle university, Newcastle upon Tyne, UK
[e]Breast Surgery, First Affiliated Hospital, Zhejiang University, Hangzhou, China
[f]The Second Affiliated Hospital of Zhejiang University School of Medicine, Zhejiang University, Hangzhou, China

ARTICLE INFO

Keywords:
Medical image segmentation
Layer normalization
Deep learning
Domain shift

ABSTRACT

The accurate segmentation of clustered microcalcifications in mammography is crucial for the diagnosis and treatment of breast cancer. Despite exhibiting expert-level accuracy, recent deep learning advancements in medical image segmentation provide insufficient contribution to practical applications, due to the domain shift resulting from differences in patient postures, individual gland density, and imaging modalities, etc. In this paper, a novel framework named MLN-net is proposed for clustered microcalcification segmentation. It can segment multi-source images using only single source images. Specifically, to rich domain distribution information, we introduce a source domain image augmentation for generating multi-source images. A structure of multiple layer normalization (LN) layers is then used to construct the segmentation network, which can be found efficient for clustered microcalcification segmentation in different domains. Additionally, a branch selection strategy is designed for measuring the similarity of the source domain data and the target domain data. To validate the proposed MLN-net, extensive analyses including ablation experiments are performed, comparison of 12 baseline methods. MLN-net enhances segmentation quality of full-field digital mammography (FFDM) and digital breast tomosynthe (DBT) images from the FFDM-DBT dataset, achieving the average Dice similarity coefficient (DSC) of 86.52% and the average Hausdorff distance (HD) of 20.49mm on the source domain DBT. And it outperforms the baseline models for the task in FFDM images from both the CBIS-DDSM and the FFDM-DBT dataset, achieving the average DSC of 50.78% and the average HD of 35.12mm on the source domain CBIS-DDSM. Extensive experiments validate the effectiveness of MLN-net in segmenting clustered microcalcifications from different domains and its segmentation accuracy surpasses state-of-the-art methods. Code will be available at https://github.com/yezanting/MLN-NET-VERSON1.

1. Introduction

Breast cancer is the most common cancer diagnosed in women, accounting for 2.3 million cases and 11.7% of all cancers worldwide Sung et al. (2021). Compelling evidences demonstrate that the establishment of a screening program can reduce breast cancer-related mortality by as much as 20% Mukama et al. (2020). Mammography is widely used for screening breast cancer, with a detection rate ranging from 80% to 90%. Clustered microcalcifications, present in 30% to 50% of breast cancer patients, constitute primary pathological features of this disease. Principal mammography modalities employed for detecting clustered microcalcifications encompass FFDM and DBT Tarver (2012); In (2019). FFDM, a 2D imaging modality, acquires two-dimensional X-ray breast images in a single view. DBT is an advanced 3D imaging technology that captures a sequence of low-dose X-ray images of the breast from varying angles, subsequently reconstructing these images into thin tomographic slices with a 1 mm separation Chong et al. (2019); Zuckerman et al. (2020). Microcalcifications manifest as bright spots in both FFDM and DBT, as shown in Fig. 1. The morphology, dimensions, and distribution of clustered microcalcifications hold a pivotal role in distinguishing benign and malignant breast cancer cases Chong et al. (2019). Compared with FFDM, DBT mitigates the challenge of overlapping structures, facilitating radiologists in the differentiation between healthy breast tissue and potential abnormalities. This ultimately improves breast cancer detection and reduce false-positive findings Conant et al. (2020); Richman et al. (2019). Nevertheless, FFDM continues to serve as the primary approach for detecting clustered microcalcifications in the majority of countries, chiefly due to the following three factors: 1) DBT entails multiple low-dose X-ray images, resulting in a higher cumulative radiation dose compared to FFDM Liu et al. (2018). 2) The data volume of DBT is a hundredfold greater than that of FFDM, leading to extended acquisition and interpretation times attributable to the larger image dataset Chong et al. (2019). 3) DBT is a more recent technology, and its adoption may vary by location and healthcare provider. Radiologists require additional training to interpret DBT images. The concern regarding radiation dose is gradually being addressed through advancements in imaging technology, leading to increased DBT utilization in recent years. In this paper, we aim to enhance the efficiency

*Corresponding authors.
yzt2861252880@gmail.com (Z. Ye); 3110102872@zju.edu.cn (B. Liu)
ORCID(s): 0000-0002-1926-4634 (K. Wang); 0009-0006-8874-8882 (Z. Ye)

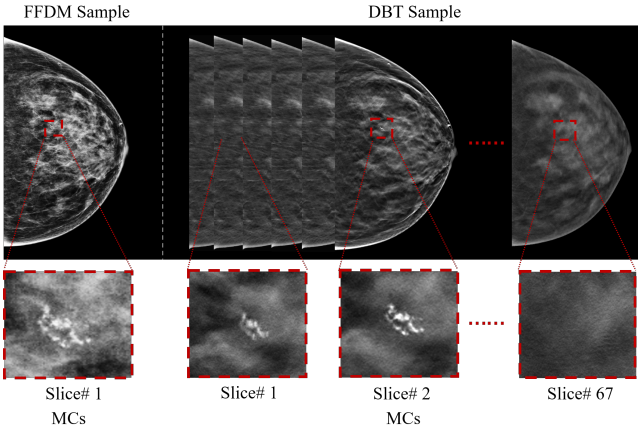

Figure 1: Examples of images from FFDM and DBT imaging modalities. The images are from the same patient and represent the same area of breast tissue. Clustered microcalcifications can be seen in both imaging modalities. Compared to FFDM, DBT performs slice photography of the lesion with a much narrower interval of 1mm, producing dozens to hundreds of images depending on individual differences.

and precision of clustered microcalcifications recognition within the FFDM modality from diverse sources, while also exploring the transfer of knowledge acquired in the FFDM modality to the realm of DBT.

Deep neural network (DNN) is widely recognized as a state-of-the-art technology for image analysis, delivering superior performance in various tasks such as image classification, semantic segmentation, and object detection Miao et al. (2023); Ibtehaz and Rahman (2020); Miao et al. (2022); Zhou et al. (2022); He et al. (2022). The sharing of a substantial volume of high-quality datasets has significantly bolstered the advancement of DNNs Miao et al. (2021); Lee et al. (2017). DNN-based computer-aided diagnosis (CAD) for clustered microcalcifications is rapidly and effectively being integrated into systems engineering. These methods use deep learning technology for automated screening of clustered microcalcifications in mammograms, reducing the workload of medical professionals and enhancing the accuracy of diagnosis. For instance, Wang and Yang (2018) proposed a context-sensitive deep learning approach that utilized the local features of clustered microcalcifications as well as the surrounding tissue background to reduce false positives in the output. Carneiro et al. (2017) proposed an automated methodology that combined information from cranio-caudal (CC) and medio-lateral oblique (MLO) views in FFDM and applied a deep learning model to extract the features. The feature was then used to assess the risks of lesions. Zheng et al. (2020) proposed a novel 3D Context-Aware Convolutional Neural Network (CNN) for improving the accuracy of clustered microcalcifications detection in DBT. The method utilized a 2D CNN to extract the intra-slice features of DBT images, while a 3D CNN was employed to extract the inter-slice features. This method effectively employed multi-fault information in DBT to reduce false positives in the output. Samala et al. (2016) reconstructed the DBT volume using a multiscale bilateral filtering regularized simultaneous algebraic reconstruction technique, and subsequently used a CNN model to extract relevant features from the reconstructed volume. DNN-based methods assume that the source (training data) and target (test data) domains, as well as the the marginal distributions of the data, are supposed to be the same Kowald et al. (2022), resulting in a close coupling of methods with data sources. While these DNN-based methods achieve high sensitivity in clustered microcalcifications detecting, they often face a reduction in performance when there is a substantial difference between the source and target domains. The phenomenon commonly referred to as the domain shift problem. Although a labelled dataset can help alleviate this problem, it can be costly for collecting such labeled medical datasets, and so these approaches are impractical.

Domain shifts frequently occur in clinical practice, presenting a significant obstacle to the effective use of DNNs. Factors such as the continuous advancement of image acquisition technology, variation in diagnostic procedures, diversity of scanners, and evolving imaging protocols may cause a decline in prediction accuracy on new data or render models obsolete due to these domain shifts Perkonigg et al. (2021). Our research focuses on identifying clustered microcalcifications in breast cancer, with a particular emphasis on countering the domain shift challenge in mammography. In mammography, gland density disparities among individuals, patient positioning, and exposure intensity variances on the same equipment contribute to intrinsic data variability. This issue becomes even more prominent when data is sourced from divergent imaging modalities (distinct imaging protocols), such as FFDM and DBT, representing extrinsic data variability.

Therefore, we propose MLN-net, which accepts single-domain data and can generalize to unseen domains. At its core, MLN-net seeks to tackle the domain shift issue, specifically when handling disparate clustered microcalcifications datasets. MLN-net has been meticulously designed to mitigate data variation impacts, enabling its application to new target datasets without necessitating retraining with labeled data. Furthermore, MLN-net is a medical image segmentation method, providing more detailed lesion information for the diagnosis and treatment of breast cancer.

Specifically, MLN-net reconstructs the original data exactly from the single-domain set to multi-domain under the Bézier curve and grayscale-inversion transformation, fostering greater data diversity and simulating potential shifts in the target domain. To perform multi-domain feature extraction, we propose a novel segmentation network with multiple LN layers. This network enhances domain information capture and reduces computation through the shared use of features among different inputs. Lastly, we propose a branch selection strategy, hinged on distance metric, for optimal LN layer selection within the segmentation network. During testing, this strategy gauges the distribution difference

between the target and source domains by computing cosine similarity, thereby selecting the best segmentation results.

In summary, our contributions are summarized as follows:

- We propose a novel framework for segmenting clustered microcalcificationsm, called MLN-net. MLN-net is capable of accurately segmenting multi-source images using only single-source images, thus effectively addressing the issue of domain shift between the source and target domains.
- We introduce a novel segmentation network with multiple LN layers to capture the features of multi-source images.
- We develop a source domain data augmentation based on Bézier curves and grayscale-inversion transformation to increase the diversity of the source domain data. Additionally, we develop a branch selection strategy to measure the similarities between the source domain data and the target domain data.
- We evaluate the proposed MLN-net on both the private FFDM-DBT dataset and the publicly available CBIS-DDSM dataset. MLN-net achieved superior performance over state-of-the-art methods. And in-depth analytical experiments demonstrate the efficacy of MLN-net.

2. Related work

2.1. Domain shift

Domain shifts arise from differences between the distribution of the training and the test data. In medical image analysis, these differences can result from different scanners, scanner generations, manufacturers, or imaging protocols during data acquisition Perkonigg et al. (2021); He et al. (2021). For example, breast cancer diagnosis involves making use of medical image information obtained from both DBT and FFDM Horvat et al. (2019); Giess et al. (2017), corresponding to different imaging protocols. Even with the same imaging protocols, shown in Fig. 7, there can be significant discrepancies between images. These image disparities are a result of the rapid evolution of mammography technology. Therefore, to ensure successful deployment of deep learning-based clustered microcalcifications identification models in the changing environment, it is crucial to develop and advance methods that consider these domain shifts.

Recently, several methods have been proposed to address the issue of domain shift, encompassing semi-supervised transfer learning (STL) Wang et al. (2021); Abuduweili et al. (2021); Zhang et al. (2021); Jakubovitz et al. (2019); Wei et al. (2019), unsupervised domain adaptation (UDA) Du et al. (2019); Dong et al. (2020); Doersch et al. (2015), and domain generalization (DG) Fan et al. (2021); Seo et al. (2020); Zhang et al. (2020); Zhou et al. (2020). Table 1 presents a comprehensive overview of the merits and demerits found in a range of related research endeavors, along with the primary technologies employed therein.

2.1.1. Semi-supervised transfer learning

STL constructs effective self-supervised mechanisms for unlabeled data to migrate the rich knowledge of the source domain to the target domain. For instance, Abuduweili et al. (2021) proposed an adaptive consistency regularization approach to leverage both pre-trained weights and unlabeled data. The adaptive consistency regularization approach consists of two complementary components: Adaptive Knowledge Consistency, which applies to the examples between the source and target models, and Adaptive Representation Consistency, which applies to the target model between labeled and unlabeled examples. Wang et al. (2021) utilized a pretrained model and a novel transfer feature learning model to extract image features, followed by proposing a deep fusion method and a selection method to combine the features from the pretrained and transfer feature learning model. This architecture enhanced the model's generalization capability.

2.1.2. Unsupervised domain adaptation

Unlike STL, UDA does not require a large amount of relevant data. UDA mitigates the distribution difference between the source and target domains to address the domain shift issue. For example, Du et al. (2019) and Dong et al. (2020) applied deep adversarial networks to decrease the feature distribution gap among data from different domains so that the model can learn the same semantic features. Another noteworthy contribution includes He et al. (2021), which presented a system of three neural networks: the task model, the autoencoder model, and the adaptor model. The task model undertaken the image analysis task, while the autoencoder and adaptor models transfigured the features of the target domain to minimize domain shift. Other methods, such as those proposed by Liu et al. (2020b) and Ma et al. (2019), adopted style transfer methods to adapt the transformation of the target domain by analyzing the differences between the source and target domains.

2.1.3. Domain generalization

STL and UDA face challenges when applied in the medical field due to privacy concerns that hinder the sharing of data between different hospitals and departments. These privacy concerns result in the target domain being unknown during the training process. Unlike STL and UDA, DG relies solely on the source domain data and builds models capable of directly generalizing to target domains, presenting a more pragmatic solution. Several DG methods have been proposed recently Zhou et al. (2022); Zhang et al. (2020); Balaji et al. (2018); Dou et al. (2019); Liu et al. (2020a); Fan et al. (2021); Segu et al. (2023); Seo et al. (2020); Zhou et al. (2020). Meta-learning, for instance, has been utilized to facilitate domain-invariant representation of multi-domain data Balaji et al. (2018); Dou et al. (2019); Liu et al. (2020a). There are also investigations that enhance the normalization process of DNNs by incorporating improved batch normalization (BN) and instance normalization (IN), thereby

Table 1
Summary of related methods to solve domain shift. We evaluated the merits and demerits of each method based on four aspects: whether it requires a large amount of relevant data, the availability of target domain data, its suitability for large-scale domain shift, and small target segmentation tasks. "-" means it cannot be determined or defined.

Method	Literature	Demerits				Main Technique
		Require a large number of relevant data	Require available target domain data	Unsuitable for large scale domain shift	Unsuitable for small target segmentation	
STL	Wang et al. (2021)	✓	✓		-	Transfer learning and Fine-tuning
	Abuduweili et al. (2021)	✓	✓		-	Transfer learning and Fine-tuning
UDA	Du et al. (2019)		✓	-	✓	Deep adversarial network
	Dong et al. (2020)		✓	-	✓	Deep adversarial network
	He et al. (2021)		✓	-	-	-
	Liu et al. (2020b)		✓	✓		Style transfer
	Ma et al. (2019)		✓	✓		Style transfer
DG	Balaji et al. (2018)			✓	✓	Meta-learning
	Dou et al. (2019)			✓	✓	Meta-learning
	Liu et al. (2020a)			✓	✓	Meta-learning
	Fan et al. (2021)				✓	Normalization strategy
	Segu et al. (2023)				✓	Normalization strategy
	Seo et al. (2020)				✓	Normalization strategy
	Zhang et al. (2020)			✓	✓	Data augmentation
	Zhou et al. (2022)			✓	✓	Data augmentation and normalization strategy
	MLN-net					Data augmentation and normalization strategy

strengthening the networks' capacity to grasp domain information Fan et al. (2021); Segu et al. (2023); Seo et al. (2020). Additionally, strategies such as specific style transfer or data augmentation, which indirectly tackle potential changes in the target domain, have been successfully applied in medical image analysis to mitigate the issue of domain shift Zhang et al. (2020); Zhou et al. (2022).

Clustered microcalcifications identification is a critical task in clinical practice Pan et al. (2022). Existing DG methods are not directly applicable to the clustered microcalcifications segmentation task, mainly for two reasons. First, segmentation of the clustered microcalcifications involves distinction between the lesion's characteristics and the background, which presents a challenging classification problem. Second, most DG methods is relatively sensitive to small shifts of domain distribution and so is unable to provide a accurate segmentation to large domain shifts. To overcome these limitations, we propose a novel segmentation method for clustered microcalcifications that is the first segmentation method with domain generalization ability for clustered microcalcifications, filling a critical gap in the field.

2.2. Medical image segmentation based on deep learning

Medical image segmentation is a crucial component of automated medical image analysis, as it allows for the extraction of essential quantitative imaging markers, which in turn improves diagnosis, personalized treatment planning, and therapy monitoring Sourati et al. (2019). With the emergence of deep learning, segmentation approaches have evolved from traditional machine learning models to deep learning methods, yielding promising results in various segmentation tasks Jiang et al. (2020).

2.2.1. Convolutional neural network

Convolutional neural network (CNN) He et al. (2016), which utilizes convolution and pooling operations to extract image features, have found widespread application in the field of medical imaging. CNNs offer significant benefits in image classification and object detection, particularly in the context of clustered microcalcifications Bekker et al. (2015); Wang and Yang (2018); Zheng et al. (2020); Carneiro et al. (2017). However, image segmentation, a pixel-level classification task, necessitates a meticulous representation of image features often compromised during down-sampling in pooling operations. Moreover, the scarcity of semantic information in shallow feature maps generated by CNNs typically results in subpar segmentation performance, especially when dealing with small targets set against complex backgrounds.

2.2.2. Vision transformer

In recent years, vision transformer methods Vaswani et al. (2017); Cao et al. (2023); Li et al. (2022), based on attention mechanism, have been introduced to exploit the correlations that exist in medical images between pixels. In contrast to CNNs, these methods excel at preserving detailed features and demonstrate superior feature representation capabilities Cao et al. (2023); Chen et al. (2021); Bougourzi et al. (2023); Gu et al. (2020). Representatively, Swin-Unet Cao et al. (2023) and Transunet Chen et al. (2021) incorporated a self-attention mechanism into the basic Unet framework Ronneberger et al. (2015), enhancing the global modeling ability and achieving excellent segmentation performance. PDAttunet Bougourzi et al. (2023), the latest work, which used the attention mechanism to segment the COVID-19 lesions. In our study, we employ Swinuet Cao et al. (2023) as the backbone of segmentation network to extract image features. Diffident from the previous work Cao et al. (2023), we introduce a multiple LN layers structure into the network, thereby improving its capability to capture multi-domain information.

3. Method

Let $D^s = \sum_{i}^{N} \{x_i, y_i\}$ represents a source domain set, where S denotes the domain label, x_i is the i_{th} image in the domain, y_i signifies the ground truth, and N is the total

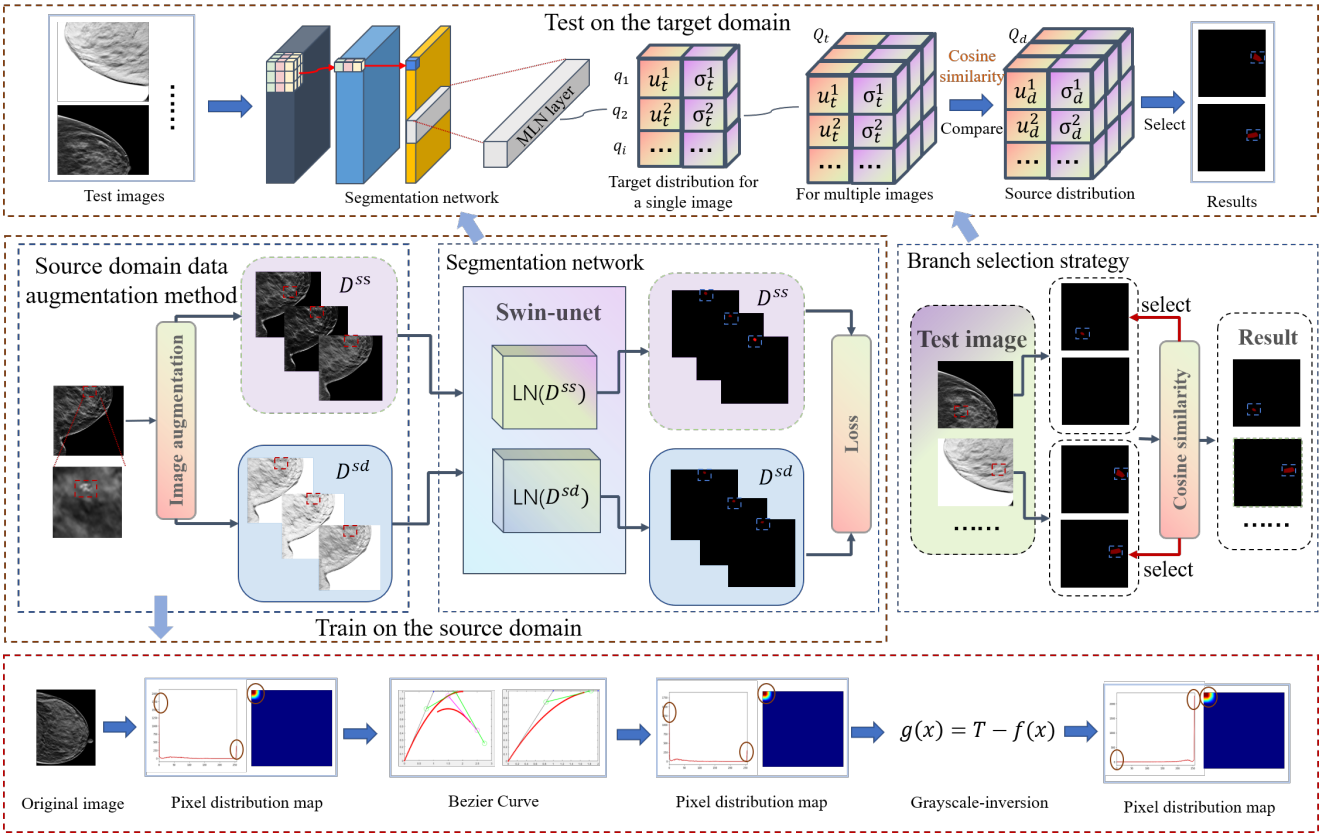

Figure 2: The schematic diagram of MLN-net framework. MLN-net utilizes source domain data for generalizable segmentation on unseen domain data. MLN-net is composed of a source domain data augmentation, a segmentation network, and a branch selection strategy. During training stage, the source domain data augmentation is utilized to augment the source domain data. The augmented data is then fed to the segmentation network with multiple LN layers, and the loss functions is used to optimize the segmentation results of different branches. During testing stage, the test data is fed into the trained segmentation network. And the branch selection strategy is adopted to choose the optimal segmentation results.

number of samples within the domain. Our primary objective is to develop a multi-source image segmentation model M_ϕ that can achieve excellent generalization performance on the source domains D^{ss} and the target domains D^{sd}. The segmentation model takes the form:

$$M_\phi(x) = y_p, (x) \in D^{ss}, D^{sd}, \tag{1}$$

$$\min(y_p - y_g) = 0. \tag{2}$$

where ϕ represents the parameters of segmentation model, x denotes the input image, y_p signifies the mask of prediction, and y_g signifies the mask of ground truth.

The overall pipeline of MLN-net is illustrated in Fig. 2. MLN-net comprises three primary modules: a source domain data augmentation, a segmentation network, and a branch selection strategy. This framework facilitates a generalizable segmentation approach that can extract features from the source domain data and segment clustered microcalcifications on unseen domains. Subsequent sections will provide a detailed explanation of MLN-net's architecture.

3.1. Source domain data augmentation

Data augmentation is frequently employed to improve accuracy in neural network recognition systems for medical data. And recent studies have suggested that data augmentation can alleviate the problem of domain shift Zhang et al. (2020); Zhou et al. (2022). Motivated by these insights, we present, as depicted in Fig. 3, a monotonic non-linear transformation function that uses the Bézier curve to adjust pixel values, acquiring the source-similar data. Furthermore, we introduce a grayscale-inversion transformation method to acquire the source-dissimilar data. Algorithm 1 elucidates the comprehensive procedure of the source domain data augmentation.

3.1.1. Source-similar data augmentation

Mammography images, such as FFDM and DBT, are typically grayscale images. As depicted in Fig. 1, the distribution, boundary, and morphology of clustered microcalcifications in both types of images share similarities. The primary distinguishing factor is the pixel mapping strategy that defines light and dark features. Inspired by this observation, we propose a straightforward source-similar data augmentation method that utilizes the Bézier curve to transform the gray distribution of images. This method enhances the use

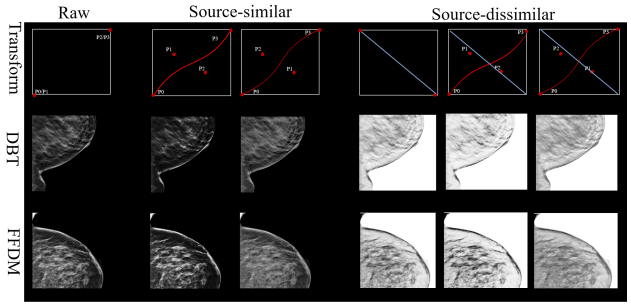

Figure 3: Examples of source-similar and source-dissimilar transformations results on the FFDM-DBT dataset. In the first row, the red line signifies the transformation of source-similar based on the Bézier curve. This curve has four critical points; P1 denotes the starting point while P4 denotes the endpoint. P2 and P3 denote the control points that determine the curvature of the Bézier curve. In contrast, the blue line signifies transformation of source-dissimilar based on grayscale-inversion. And the second and the third row are the images after the transformations

of the source domain data and facilitates the generalization of the model to unseen target domains. The Bézier curve has a starting point, an end point, and two control points. Its mathematical expression can be represented as:

$$B(k) = \sum_{i=0}^{N} \binom{N}{i} P_i (1-k)^{N-k} k^i, N = 3, k \in [p_{low}, p_{high}], \tag{3}$$

where k is a fractional value along the line's length. In our study, we normalize the pixel values of data to the range [0,1]. Therefore, the total length of the line is fixed at 1, divided into 10,000 segments. $P_{start} = (p_{low}, p_{low})$ and $P_{end} = (p_{high}, p_{high})$ are the start and end points, which are set to $(0, 0)$ and $(1, 1)$, respectively, based on the pixel distribution value range. A series of P_c represents the control point with values randomly generated from the interval $[p_{low}, p_{high}]$. To augment the data, we employ two sets of curve control points. A comprehensive discussion of these parameter settings will be provided in the following ablation experiment.

Algorithm 1 elucidates the comprehensive procedure of our augmentation pipeline, whereas the aforementioned source-similar data augmentation is expounded in lines 1-19. The process of source-similar data augmentation unfolds as follows: Initially, we initialize the parameters and normalize the input images. Subsequently, two sets of curve control points P_c are randomly generated from the interval $[p_{low}, p_{high}]$. Next, the pixel distribution of input images is adjusted using Bézier curves to generate source-similar data, as detailed in lines 4-19 of Algorithm 1.

3.1.2. Source-dissimilar data augmentation

In mammography images, lesions appear in the form of small white spots with black background, which results in decreased reliability in lesion identification. Besides, different imaging modalities apply varied pixel mapping approaches, further complicating the detection of pathological features. Standard DNN models that rely solely on superficial features may perform inadequately when presented with domain shifts. Because of the close similarity between the source-similar data, it is difficult to identify the corresponding internal feature mapping on domain shift tasks. Data augmentation method can therefore be used to give rise to variations in source-dissimilar data, in order to obtain a good generalization performance.

The grayscale-inversion transformation, as an image augmentation method, is commonly used in medical image processing to enhance the visibility of lesion areas. In our study, we introduce a novel use-case, employing the grayscale-inversion transformation, as a data augmentation method, to encourage an adaptive model. The training set is augmented using replicas of the training patterns, transformed according to the desired invariances. Specifically, anti-gray images, defined as source-dissimilar data, is introduced to allow a more complex target domain in the training dataset. This greatly reduces the dependence of the classification performance on notable features, which provides a framework for feature extraction with broad

Algorithm 1: Source Domain Data Augmentation

Input: Training data (x), P_{start}, P_{end}, two sets of curve control points P_c
Output: Augmented data x^{aug}
1: Initialize the parameters
2: $x^{nor} \leftarrow$ Normalize(x)
3: Randomly generate two sets of curve control points P_c from the interval $[p_{low}, p_{high}]$
4: Input P_{start}, P_{end} and two sets of curve control points P_c
5: $N \leftarrow$ len(P) ▷ Calculate total number of P
6: for p in P do
7: $x_{points} \leftarrow$ p[0]
8: $y_{points} \leftarrow$ p[1]
9: end
10: for $Times \leftarrow$ 0 to 10000 do
11: $k[Times] \leftarrow Times/10000$
12:end
13:for $i \leftarrow$ 0 to N do
14: $L_{array} \leftarrow B(k)$ ▷ See Eqn.3
15:end
16:$x^{vals} \leftarrow x_{points} \cdot L_{array}$
17:$y^{vals} \leftarrow y_{points} \cdot L_{array}$
18:$x^{sort} \leftarrow$ sort(x^{vals})
19:$y^{sort} \leftarrow$ sort(y^{vals}) ▷ Sort in ascending order
20:$x^{similar} \leftarrow$ Interp$(x^{nor}, x^{sort}, y^{sort})$ ▷ Interpolation
21:$x^{dissimilar} \leftarrow p_{high} - x^{nor}$ ▷ See Eqn.4
22:$x^{aug} \leftarrow x^{similar} \odot x^{dissimilar}$
23:**Return** x^{aug}
▷ $\odot$ indicates mutually exclusive events (which is inactive in the case of monotonic non-linear transformation)

applicability. The mathematical expression of the grayscale-inversion transformation is presented as follows:

$$g(x, y) = T - f(x, y). \tag{4}$$

where T is the maximum pixel value of the image, f is the pixel value of the current position. x, y are the coordinate positions, respectively.

3.2. Segmentation network

3.2.1. Backbone of segmentation network

In order to segment clustered microcalcifications across diverse domains, a novel segmentation network with multiple LN layers is introduced. This network utilizes Swin-Unet Cao et al. (2023) as its backbone, with the self-attention mechanism for advanced feature representation learning. As depicted in Fig. 4, the Swin-Unet block is the heart of the Swin-Unet, primarily comprises of the window-based multi-head self-attention (W-MSA) and shifted window-based multi-head self-attention (SW-MSA) module. A window partitioning approach can therefore be proposed, which applies multi-head self-attention modules to two consecutive blocks. And successive Swin-Unet Transformer blocks based on this approach can be represented as follows:

$$\tilde{x}^t = W - MSA(LN(x^t)) + x^{t-1}, \tag{5}$$

$$x^t = MLP(LN(\tilde{x}^t)) + \tilde{x}^t, \tag{6}$$

$$\tilde{x}^{t+1} = SW - MSA(LN(x^t)) + x^t, \tag{7}$$

$$x^{t+1} = MLP(LN(\tilde{x}^{t+1})) + \tilde{x}^{t+1}. \tag{8}$$

where x^t is the output of layer t. LN represents the LN layer.

The self-attention mechanism of Swin-Unet can be viewed as performing a feature extraction. And information from such features can then be used to address long-term dependency issues prevalent in CNN-based methods Zhang et al. (2022). Unlike conventional self-attention based methods, Swin-Unet can capture the correlations between each region in the image and reduce computational overhead Cao et al. (2023).

3.2.2. Multiple LN layers structure

LN, a data normalization technique, can reduce the impact of internal covariance transformation and resolve the problem of vanishing and exploding gradients Ba et al. (2016). Unlike the commonly used BN method Ioffe and Szegedy (2015), LN is insensitive to the input data batch size. This makes it an efficient normalization method in terms of self-attention based architectures. The mathematical expression of LN can be written as:

$$\mathrm{u}^l = \frac{1}{n}\sum_{i=1}^{n} h_i^l, \tag{9}$$

$$\sigma^l = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(h_i^l - \mathrm{u}^l)^2}, \tag{10}$$

$$L^t = \frac{\gamma}{\sigma^l} \odot (h^t - \mathrm{u}^l) + \omega. \tag{11}$$

where h_i^l represents the i element of the l layer, n represents the total number of elements at l layer, u^l, σ^l are mean and standard deviation; γ, ω are parameters for scaling and translation.

Data standardization is applied in LN to mitigate the variability in input data. However capturing domain distribution information from multiple domains poses a significant challenge. The deficiency of domain distribution information results in the overall loss incurred in decision making, which can lead to poor generalization and robustness. To resolve this issue, we introduce a segmentation network with multiple LN layers to extract the required distribution information of varying domains. The segmentation network utilizes the same backbone network parameters but different normalization strategies, which takes the form:

$$MLN(d) = \frac{\gamma_d}{\sqrt{\sigma_d^2 + \Delta}}(x - \mathrm{u}_d) + \omega_d. \tag{12}$$

where d is the label of domain, u, σ are mean and standard deviation, Δ is a small constant value to avoid $\sqrt{\sigma_d^2 + \Delta} = 0$, γ_d, ω_d are parameters for scaling and translation of d domain.

During the training process, domain-specific information can be estimated by recording the parameters of the normalization layer for each domain at the final iteration. Evaluation of the best branch is straightforward as each normalization layer only involves the u and σ values, which are recorded in training and test process.

3.2.3. Loss function

MLN-net is modeled with multiple LN layers corresponding to a multi-branch structure. To evaluate the performance of all branches, we can formalize such issue through the introduction of a Dice loss, which is a overall measure of loss incurred in each branch. In this case, the overall loss can be written:

$$f_{dice}^k = -\frac{2}{N}\sum_{n\in N}\frac{\sum_i c_{i,n} z_{i,n}}{\sum_i c_{i,n} + \sum_i z_{i,n}}, \tag{13}$$

$$L = \sum_{b=1}^{k} f_{dice}^b. \tag{14}$$

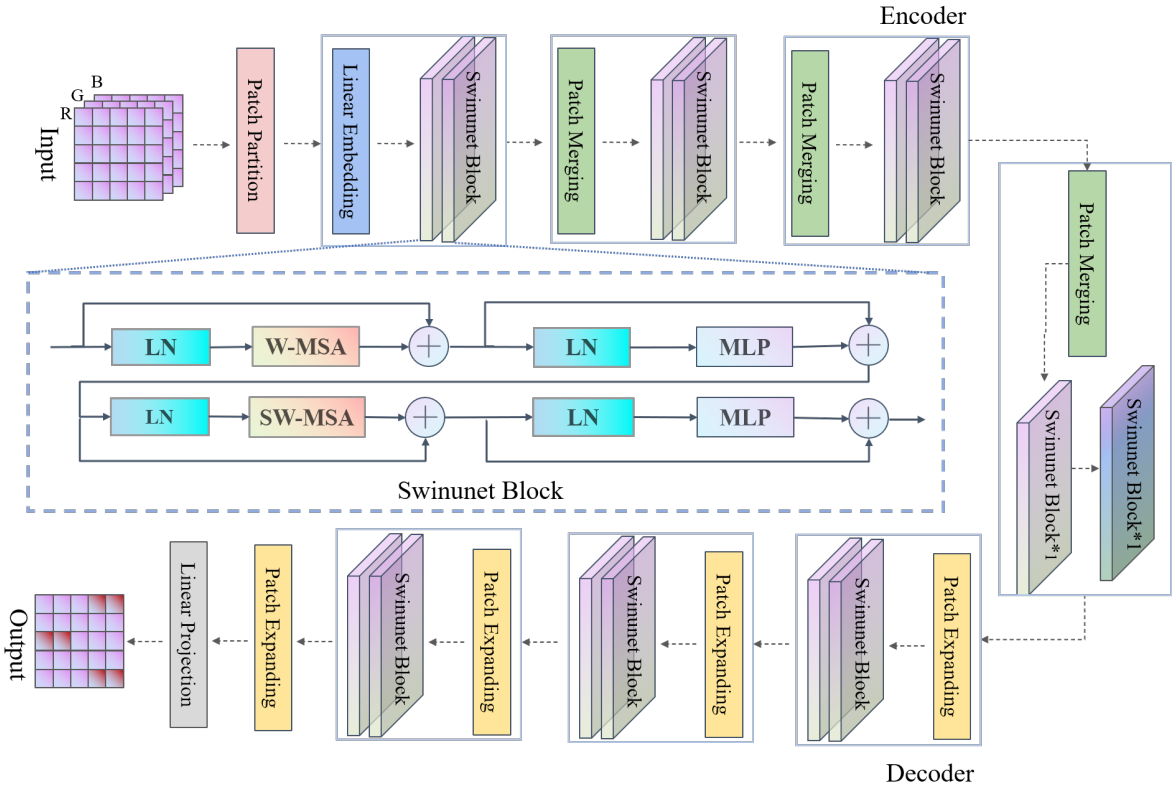

Figure 4: The composition structure of the backbone network and the Swin-Unet blocks in the Swin-Unet model.

where k is the k_{th} branch, N is the total number of classification categories, $c_{i,n}$ is the softmax output that classifies the pixel i into the class n, and $z_{i,n}$ is the ground truth of pixel i.

3.3. Branch selection strategy

The normalization process in segmentation network is expressed in terms of multiple LN layers, which leads to multiple outputs. And we can provide a much more general view of domain distribution information by giving a probabilistic interpretation to the multiple segmentation outputs Segu et al. (2023). A branch selection strategy grounded in distance metrics is devised, which calculates the distance between the mean and standard deviation of both the target and source domains. Thus the domain distribution information is given by:

$$Q_d = [q_d^1, q_d^2, ..., q_d^l] = [(u_d^1, \sigma_d^1), (u_d^2, \sigma_d^2), ..., (u_d^l, \sigma_d^l)], \tag{15}$$

$$Q_t = [q_t^1, q_t^2, ..., q_t^l] = [(u_t^1, \sigma_t^1), (u_t^2, \sigma_t^2), ..., (u_t^l, \sigma_t^l)]. \tag{16}$$

where d represents the label of source domain, t represents the label of target domain, and l represents the l_{th} LN layer in MLN-net.

In practice, it is often feasible to use the Euclidean distance metric for similarity evaluation. In the branch selection strategy, we replace the traditional Euclidean distance with cosine distance to measure similarity between Q_d and Q_t, as cosine distance metric has superiority in quantifying similarity of high-dimensional space vectors Strehl et al. (2000). The corresponding strategy is given by:

$$C(Q_d^i, Q_t^i) = \frac{q_d^i \cdot q_t^i}{\left|q_d^i\right| \left|q_t^i\right|}, \tag{17}$$

$$\text{Distance}(Q_d, Q_t) = \sum_{i=1}^{l} C(Q_d^i, Q_t^i), \tag{18}$$

$$S = \arg\min(\text{Distance}(Q_d, Q_t)). \tag{19}$$

where d represents the source domain, t represents the target domain, i represents the i_{th} LN layer of MLN-net. C is cosine similarity.

4. Experiments

4.1. Experiments setup

4.1.1. Data and pre-processing

The private dataset FFDM-DBT and the public dataset CBIS-DDSM Lee et al. (2017) are used to validate our methodology for segmenting clustered microcalcifications. The FFDM-DBT dataset consists of DBT data from 80

patients and FFDM data from 420 patients, each with corresponding pixel-level labels, sourced from the Breast Surgery department at the First Affiliated Hospital, Zhejiang University. DBT data, each with a thickness of 1mm, incorporate double breast tomography images for every patient. The image count ranges from 70 to 200 per patient, maintaining a resolution of 1996×2457. Each patient's FFDM data includes four views: CC and MLO images of the left and right breasts, all at the same resolution of 1996×2457. Experienced radiologists manually annotated the labels for both FFDM and DBT data. The image preprocessing stage applies a 2×2 equal ratio segmentation on FFDM and DBT data to exclude the background blocks, and subsequently resamples the images to a 512×512 resolution, readying them for the source domain data augmentation. For this study, we randomly selected 80% of the data for the training set and reserved the remaining 20% for the test set.

The CBIS-DDSM dataset includes 735 original breast images, each paired with corresponding pixel-level labels, procured from Massachusetts General Hospital, the University of South Florida, and Sandia National Laboratories. Similar to the FFDM-DBT dataset, data preprocessing procedures, such as image augmentation and resampling, are performed on the CBIS-DDSM dataset, thereby tailoring the data for our study.

4.1.2. Evaluation metrics

To evaluate the performance of MLN-net, we used five evaluation metrics: True Positive Rate (TPR), Precision (Pr), DSC, HD, and Average Surface Distance (ASD). TPR gauges the ability of MLN-net to detect lesions. Pr appraises its proficiency in identifying positive samples, whereas DSC assesses the extent of overlap between the predicted and actual annotations. The spatial overlap index for these three metrics (TPR, Pr, and DSC) ranges from 0 to 1, a higher value signifying superior segmentation performance. Conversely, HD and ASD evaluate the segmentation error, with lower values denoting superior segmentation performance. The metrics are defined as follows:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{20}$$

$$\mathrm{Pr} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{21}$$

$$\mathrm{DSC} = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}, \tag{22}$$

$$\mathrm{ASD} = \frac{1}{n} \sum_{p \in P} \min_{g \in G} ||p - g||, \tag{23}$$

$$\mathrm{HD} = \max \left(\sup_{p \in P} \inf_{g \in G} ||p - g||, \sup_{g \in G} \inf_{p \in P} ||g - p|| \right). \tag{24}$$

where TP is the number of true positives, TN is the number of true negative and FP is the number of false positives. P and G are the surface voxel set of predicted segmentation results and ground truth, respectively. s and g are an arbitrary voxel in P and G, and n is the total number of S and G elements. $\|\cdot\|$ denotes the shortest Euclidean distance.

4.1.3. Implementation details

The implementation of the proposed MLN-net is based on Pytorch. The experimentation was conducted on a system operating with Windows 10 and equipped with an NVIDIA GeForce RTX 3080 graphics card, possessing 10GB memory. During training, the ADAM optimizer was employed with a batch size of 4, momentum set at 0.9, weight decay at 0.001, and a maximum of 100 epochs.

4.2. Experiment I: Experimental results on the FFDM-DBT dataset

In this section, the performance of the proposed MLN-net is assessed by taking experiments with the FFDM-DBT dataset, as shown in Table 2. The results are segmented into three sections: P1 exhibits the performance of four basic segmentation networks; P2 exhibits the performance of four state-of-the-art methods for clustered microcalcifications recognition; P3 shows the performance of four state-of-the-art DG methods. Additionally, Fig. 5 illustrates the segmentation results of MLN-net and other baseline methods on the target domain DBT. And Fig. 6 visualizes the segmentation results of MLN-net alongside the physician-labeled gold standard for four comprehensive cases on the target domain FFDM, each case comprising four FFDM images from distinct imaging angles. A more detailed discussion of these results will follow.

4.2.1. Comparison with the basic segmentation methods

Table 2 (P1) shows the results of four basic medical image segmentation methods. Among these methods, both Unet Ronneberger et al. (2015) and Resnet He et al. (2016) are fully convolutional neural network (FCNN) based methods. On the source domain FFDM, Unet and Resnet achieve TPR of 48.52% and 46.45%, as well as DSC of 37.32% and 37.25%, respectively. Resnet exhibites a 3.39% improvement in Pr over Unet. Since both Unet and Resnet have similar feature extraction processes, they yield comparable results. M-net Fu et al. (2018) and Unet employ a U-shaped network, but M-net introduces a novel segmentation mode with multi-scale input and side output. M-net constructs multi-level receptive fields and develops multi-distance dependence relationships to enhance predictive performance. The use of diverse scale receptive fields is beneficial in addressing the issue of image feature extraction. Furthermore, M-net utilizes side-output layers and multi-label loss functions to capture both global and local characteristics, respectively. These proposed loss functions establish an information correlation between local lesions and global images, enhancing

Table 2
Segmentation performance comparison with twelve baseline methods on the FFDM-DBT dataset. The baseline methods include P1: The basic segmentation methods, P2: The state-of-the-art methods for recognizing clustered microcalcifications and P3: The state-of-the-art DG methods. TPR, Pr, DSC, HD and ASD are used to evaluate the performance of these methods (best result is in bold for each column). Target DBT (Source FFDM): The models are trained on the domain FFDM and tested on the domain DBT, Target FFDM (Source DBT): The models are trained on the domain DBT and tested on the domain FFDM.

		Target DBT (Source FFDM)					Target FFDM (Source DBT)				
	Method	TPR (%)	Pr (%)	DSC (%)	HD (mm)	ASD (mm)	TPR (%)	Pr (%)	DSC (%)	HD (mm)	ASD (mm)
P1	Unet	48.52	35.77	37.32	51.97	20.65	52.32	36.74	43.63	44.64	21.91
	Resnet	46.45	39.16	37.25	42.05	16.76	46.07	47.46	46.98	35.32	17.06
	M-net	57.13	60.22	55.56	30.47	13.10	52.71	65.97	59.41	25.98	13.76
	Swin-Unet	50.09	46.41	46.97	29.17	13.78	53.65	51.80	57.07	26.21	13.74
P2	CS-net	65.42	68.81	66.28	28.86	10.99	67.21	69.65	67.72	28.12	11.08
	Iadml	52.24	55.82	47.67	67.91	29.64	48.07	50.32	47.64	56.42	27.03
	Musn	57.19	55.62	55.32	45.18	22.63	61.27	59.74	58.49	39.98	19.12
	CA-net	71.15	72.40	77.92	26.50	8.71	73.06	77.59	78.06	22.03	9.59
P3	BigAug	70.91	79.66	76.96	29.15	12.01	71.70	76.43	74.26	26.74	11.85
	Dofe	65.58	59.96	59.77	37.47	18.14	67.57	66.49	66.58	34.40	16.01
	Feddg	73.59	**82.74**	75.33	31.48	12.97	77.11	79.70	77.79	26.24	9.94
	Sadn	**79.36**	81.97	74.95	27.78	9.20	83.06	78.84	**87.25**	23.80	7.67
	MLN-net	78.43	81.02	**78.91**	**23.32**	**6.72**	**85.49**	**88.75**	86.52	**20.49**	**5.96**

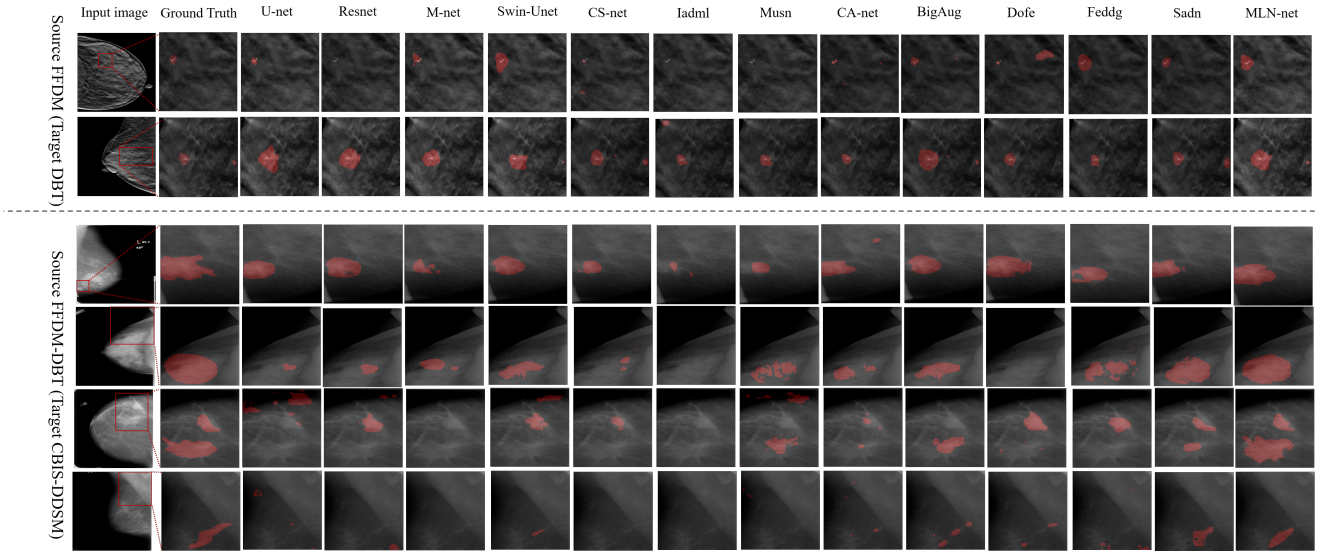

Figure 5: Clustered microcalcifications segmentation results of different methods. In the input and ground truth images, the red markings indicate the lesions annotated by doctors. The rest of the red markings represent the lesions identified by the different methods.

the segmentation power of unseen and variable target images. And on the source domain FFDM, M-net significantly enhances TPR, Pr and DSC by 8.61%, 24.45% and 18.24%, respectively, compared to Unet, confirming their beneficial effect on domain shift problem. However, Unet, Resnet, and M-net exhibited a low overall accuracy due to the FCNN's limited ability to learn explicit global and long-term semantic information interactions. This constraint directly affects their cross-center generalization ability. Moreover, FCNN's pooling operation can cause the loss of detailed features, leading to low recognition rates of small and discrete clustered microcalcifications. Swin-Unet Cao et al. (2023) leverages a U-like segmentation network that employs local window self-attention instead of convolution and pooling operations. And it employs fixed window self-attention and window self-attention sliding strategy, which results in deep feature learning while reducing computational cost. The effectiveness of Swin-Unet has proven in clustered microcalcifications segmentation tasks, achieving TPR of 53.65% , Pr of 51.80% , and DSC of 57.07% on the source domain DBT.

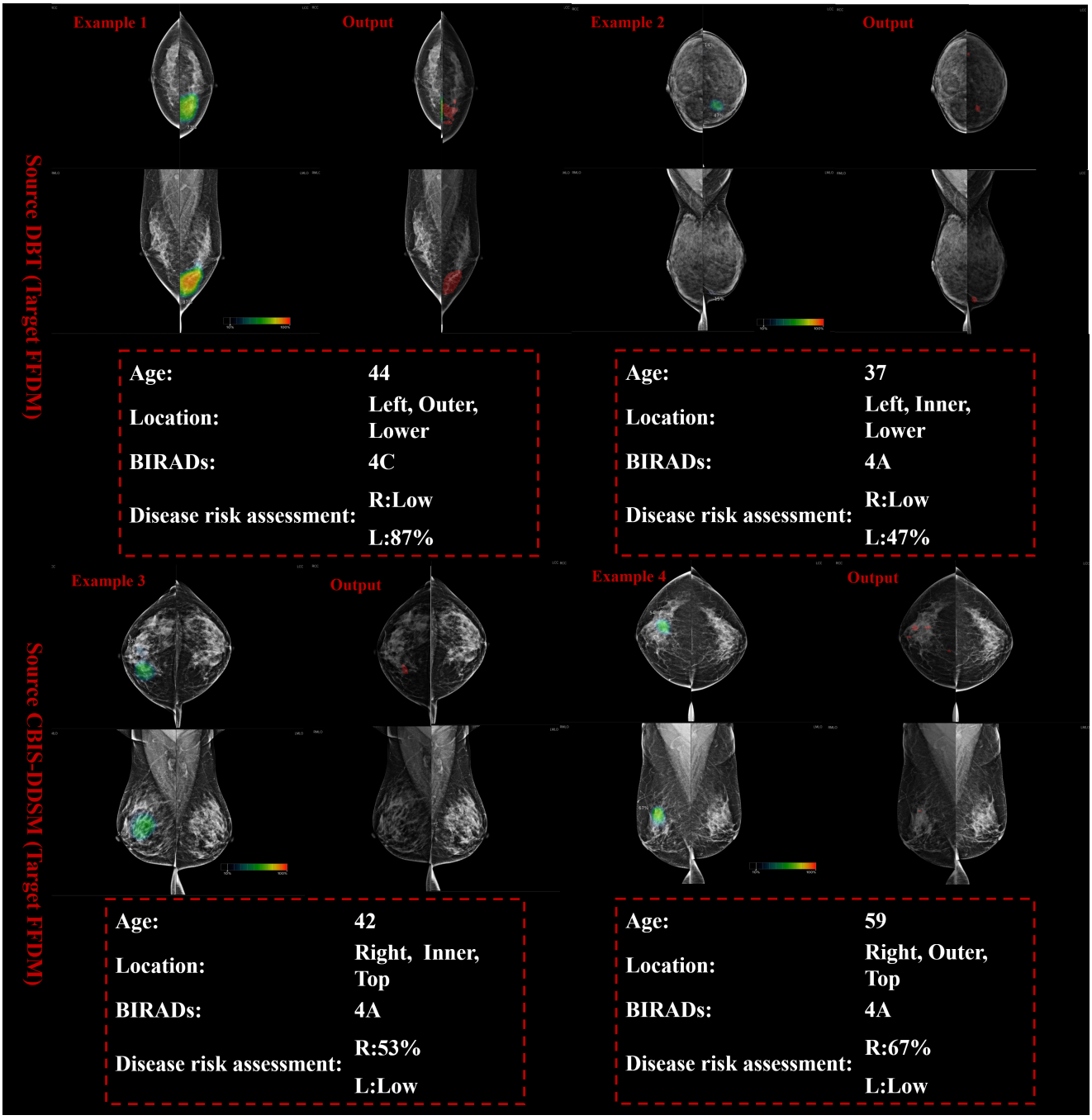

Figure 6: Example cases of MLN-net result on the target domain FFDM. This figure contains four cases, which trained on the different source domains. The left part of each case is the gold standard marked by doctors, and the right part is the output result of MLN-net. This figure shows the segmentation effectiveness of MLN-net on a complete case.

Swin-Unet is specifically designed for solving lesion segmentation within a single domain. Compared to CNNs, Swin-Unet demonstrates potential in addressing domain shift issues, which can be attributed to its ability to incorporate long-term semantic information interactions Kirillov et al. (2023); Butoi et al. (2023). In addition, every approach demonstrates appreciable performance enhancement when utilizing DBT as the source domain, in comparison to FFDM. This enhancement can likely be attributed to the wide array of tomographic images encompassed in the DBT data. These data are notably more comprehensive and encompass a broader diversity of lesion shapes and contours.

4.2.2. Comparison with the state-of-the-art methods for recognizing Clustered microcalcifications

In this section, MLN-net is contrasted with the state-of-the-art methods for recognizing clustered microcalcifications Hossain (2022); Zheng et al. (2020); Wang and Yang (2018); Wichakam et al. (2018). And the quantitative results are presented in Table 2 (P2). CS-net Wang and Yang (2018)

and Musn Hossain (2022) are developed on FFDM data. CS-net employs two parallel subnetworks to detect lesions in images of varying resolutions, effectively preserving intricate details. In our experiment, we selected images with 512x512 and 1024x1024 resolutions as CS-net's input. On the source domain FFDM, CS-net achieves TPR of 65.42% and DSC of 66.28%. These results underscore the potential advantages of multi-scale inputs in capturing domain-specific information, a finding consistent with that of M-net. Musn augments FFDM data via a local Laplacian filter and C-means clustering algorithm to select the region of interest, followed by a modified Unet for segmentation. However, Musn's excessive emphasis on local image information limits its ability to process global data, which produces a significant reduction in generalization, resulting in the subpar DSC of 45.18% on the source domain FFDM. Iadml Wichakam et al. (2018) and CA-net Zheng et al. (2020), developed on DBT data, use CNN architectures to detect clustered microcalcifications. Iadml merely uses four convolutional layers for feature extraction, which consequently leads to a potential loss of deep semantic features. In comparison, CA-net constructs a step-wise screening detection model, drawing upon a deeper CNN architecture and a more intricate detection process. It achieves DSC of 77.92% on the source domain FFDM even domain shift occur. In our implementation of CA-net, we replicated 2D FFDM and DBT images to bestow a third dimension, as CA-net applies 3D convolution for image feature extraction. This transformation, which can be perceived as a form of data augmentation, could potentially improve CA-Net's performance, possibly leading to an overestimation of its capabilities..

Current deep learning methods applied in clustered microcalcifications detection are originally designed for data on the same domain, thereby overlooking the challenge of domain shift. The swift advancement of medical imaging technology, which leads to data transformation, has profoundly curtailed the applicability of these techniques. To remedy this shortcoming, we propose MLN-net: A recognition method for clustered microcalcifications boasting cross-domain generalization capabilities. This method alleviates the intrinsic dependence of deep learning on substantial data, and holds the potential to surmount the data barriers spanning diverse imaging devices.

4.2.3. Comparison with the state-of-the-art DG methods

In this section, four state-of-the-art DG methods are chosen for comparision, including BigAug Zhang et al. (2020), Dofe Wang et al. (2020), Feddg Liu et al. (2021) and Sadn Zhou et al. (2022). BigAug, using a deep stacked transformation approach, emulates domain shifts for specific medical imaging modalities by augmenting data on a solitary source domain. Notably, on the source domain FFDM, BigAug achieves DSC, HD, and ASD of 74.26%, 26.74mm, and 11.85mm, respectively, underscoring the potency of simple data augmentation in addressing domain shift issues. Dofe, a domain-invariant feature learning method, leverages multi-source domain knowledge via the proposed domain knowledge pool to improve the generalization ability. To ensure a fair comparison between MLN-net and Dofe, we used both source-similar and source-dissimilar images as inputs to Dofe. However, Dofe's segmentation performance proves subpar, achieving DSC of 59.77% and HD of 18.14mm on the source domain FFDM. This diminished performance may result from Dofe's design specificity for Fundus image segmentation tasks, where its extensive prior knowledge may obstruct generalization to different segmentation tasks. Feddg, a meta-learning-based approach, demonstrates performance comparable to BigAug. Unlike BigAug, Dofe, and Feddg, which focus on learning or maintaining domain-invariant information, Sadn combines data augmentation methods and utilizes information from similar domains to improve generalization. This results in an improvement in the DSC by approximately 10% compared to Feddg on the source domain DBT. The success of Sadn may arise from recognizing similarities among various medical imaging technologies, where the variance is considerably less than that between different imaging technologies used for natural images.

Inspired by Sadn, MLN-net combines the multiple LN layers structure with the branch selection strategy and thereby selecting the optimal network branch. With the introduction of a self-attention mechanism, we can obtain a significant further improvement in feature extraction. Experimental results show that MLN-net gives better performance on evaluation metrics and outperforms other approaches in HD and ASD by a significant margin.

4.3. Experiment II: Experimental results on the CBIS-DDSM and the FFDM-DBT dataset

In Experiment I, the efficacy of MLN-net is validated on the FFDM and DBT images obtained from the hospital using different imaging techniques. As shown in Fig. 1, the data retain a significant degree of relevance despite variations in image acquisition. To further verify the robustness of MLN-net, we select the FFDM data from both the FFDM-DBT dataset and the CBIS-DDSM dataset. The chosen data come from varied hospitals and equipment, resulting in a notable disparity in pixel distribution, as illustrated in Fig. 7. The quantitative results are presented in Table. 3. And Fig. 5 shows the segmentation results of four cases on the target domain CBIS-DDSM. Further details on the results are discussed below.

4.3.1. Comparison with the basic segmentation methods

The results from four basic medical image segmentation methods are presented in Table. 3 (P1). The performance of these methods has significantly declined. For instance, M-net, which exhibited superior segmentation performance in Experiment I, only achieves a TPR of 9.87% and a DSC of 12.32% in this experimental setting. The method's multi-scale input and output strategy has proved effective in segmentation in the presence of minor domain shifts. However,

Table 3
Segmentation performance comparison with twelve baseline methods on the FFDM-DBT and the CBIS-DDSM dataset. The baseline methods include P1: The basic segmentation methods, P2: The state-of-the-art methods for recognizing clustered microcalcifications and P3: The state-of-the-art DG methods. TPR, Pr, DSC, HD and ASD are used to evaluate the performance of these methods (best result is in bold for each column). Target FFDM-DBT (Source CBIS-DDSM): The models are trained on the domain CBIS-DDSM and tested on the domain FFDM-DBT, Target CBIS-DDSM (Source FFDM-DBT): The models are trained on the domain FFDM-DBT and tested on the domain CBIS-DDSM.

		Target FFDM-DBT (Source CBIS-DDSM)					Target CBIS-DDSM (Source FFDM-DBT)				
	Method	TPR (%)	Pr (%)	DSC (%)	HD (mm)	ASD (mm)	TPR (%)	Pr (%)	DSC (%)	HD (mm)	ASD (mm)
P1	Unet	27.62	29.39	26.97	50.92	36.69	22.14	24.08	19.92	52.51	38.16
	Resnet	24.76	27.15	27.36	58.36	47.05	19.54	22.28	21.87	50.33	38.80
	M-net	9.87	13.19	12.32	69.90	44.18	7.67	8.91	6.72	71.77	47.05
	Swin-Unet	33.39	34.29	35.10	31.25	34.75	37.81	30.15	29.97	46.04	30.58
P2	CS-net	28.71	26.94	23.19	45.16	32.63	24.45	19.80	20.07	51.37	35.87
	Iadml	7.61	10.07	8.10	77.48	52.66	8.23	8.68	6.29	81.78	59.15
	Musn	16.48	20.73	21.01	55.50	41.36	20.12	22.77	18.56	55.14	40.70
	CA-net	35.13	27.93	33.67	39.97	32.10	30.16	26.76	25.82	56.18	35.75
P3	BigAug	55.14	**56.39**	44.51	**33.64**	22.58	41.93	43.17	40.48	39.99	25.03
	Dofe	32.76	34.83	29.48	55.79	31.75	37.16	33.87	31.05	46.74	28.10
	Feddg	43.08	41.32	37.97	48.95	29.84	37.15	37.32	38.82	46.04	26.99
	Sadn	51.75	49.92	45.16	41.04	26.11	39.49	41.60	42.35	48.00	26.72
	MLN-net	**56.85**	54.57	**50.78**	35.12	**20.33**	**48.35**	**49.41**	**45.76**	**36.74**	**23.02**

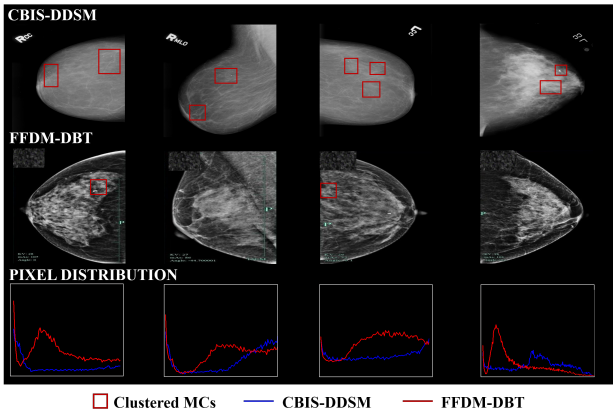

Figure 7: Example cases on the FFDM-DBT and the CBIS-DDSM dataset. Eight images from the CBIS-DDSM and the FFDM-DBT dataset are selected to perform a comparative analysis of data variations. This comparison reveals the differences in the pixel mapping approaches used by the two datasets. The third row of this figure presents a quantitative analysis of pixel values for the two types of data. Notably, we excluded from statistical analysis the points with pixel values of 0 and 1, which comprise the background. Additionally, for safeguarding patient privacy, confidential patient information is masked with grey regions on the FFDM-DBT dataset.

as the distribution differences in the data domain increase further, the method appears to excessively focus on detailed image features, thereby overlooking the global distribution information. This focus results in an inability to detect clustered microcalcifications. Unet gets DSC of 19.92% and HD of 52.51mm on the source domain FFDM-DBT. The DSC and HD decreased by about 25% and 10mm compared to the results of Table. 2. Similarly, Swin-Unet demonstrates a significant decrease in segmentation accuracy compared to Experiment I. These results suggest that the robustness and generalization performance of these methods is lacking. Consequently, their generalization quality across all target domains cannot be guaranteed when the domain shift scale is substantial.

4.3.2. Comparison with the state-of-the-art methods for recognizing clustered microcalcifications

The Table. 3 (P2) shows the result of state-of-the-art methods for recognizing clustered microcalcifications. Compared with the four basic segmentation methods, the segmentation accuracy of these methods has a significant decline. For instance, CA-net achieves DSC of 33.67% on the source domain CBIS-DDSM, indicating a decline of approximately 40% compared to Table. 2. Moreover, it is noteworthy that Iadml almost loses its ability of clustered microcalcifications segmentation, with DSC of 8.1%, HD of 77.48mm, and ASD of 52.66mm on the source domain CBIS-DDSM. The poor segmentation effect of Iadml is also revealed in the four examples in Fig. 5. In summary, the overall performance degradation of segmentation involves two key factors. First, the CBIS-DDSM dataset demonstrates limitations in effectively characterizing clustered microcalcification lesions, because of the early technological limitations. Second, as shown in Fig. 7, the disparity between the CBIS-DDSM and the FFDM-DBT datasets is greater than that between FFDM and DBT images within the FFDM-DBT dataset. These four current advanced methods for recognizing clustered microcalcifications lack robustness

in handling variations in target distribution. As a result, it can result in poor performance, especially when the dataset exhibits significant shifts in distribution.

4.3.3. Comparison with the state-of-the-art DG methods

The Table. 3 (P3) presents the results of the state-of-the-art DG methods. The results reveal that the DG methods improve the generalization performance over the methods of P2 and P1, which is largely due to their regularization effect on local learning enabling them to extract general feature representations. For example, BigAug achieves a segmentation accuracy surpassed only by the proposed MLN-net, with a DSC of 44.51%, HD of 33.64mm and ASD of 22.58mm on the source domain CBIS-DDSM, further confirming the positive impact of data augmentation. Similarly, Sadn registers commendable segmentation accuracy. As anticipated, MLN-net outperforms all other methods in terms of average results of DSC, HD, and ASD by significant margins. Specifically, on the source domain CBIS-DDSM, MLN-net achieves the highest average DSC of 50.78% and ASD of 20.33mm, which are 5.62% and 2.25mm higher than the suboptimal algorithm, respectively. And as shown in Fig. 5, the proposed MLN-net accurately segments the structure and delineates the boundary in images with unknown distributions, while other DG methods sometimes fall short. These results suggest that MLN-net exhibits superior generalization performance, even under challenging circumstances characterized by significant deviations between the source and target domains.

5. Discussion

5.1. Ablation study

MLN-net comprises three modules: the source domain data augmentation, the segmentation network, and the branch selection strategy. To demonstrate the effectiveness of each module, ablation studies were presented on the FFDM-DBT dataset. And the corresponding experiments with different configurations are given in TABLE 4, providing quantitative insights into the performance of each module.

- Basenet: The standard Swin-Unet.
- Basenet-S: The Basenet combined with the source domain data augmentation.
- MLN-net-R: The Basenet combined with the source domain data augmentation and the multiple LN layers, and without the branch selection strategy.
- Unet-R: The standard Unet combined with the source domain data augmentation and the multiple LN layers, and without the branch selection strategy.
- MLN-net-E: The Basenet combined with the source domain data augmentation, the multiple LN layers and the branch selection strategy based on Euclidean distance.
- Unet-E: The standard Unet combined with the source domain data augmentation, the multiple LN layers and the branch selection strategy based on Euclidean distance.
- MLN-net: The Basenet combined with the source domain data augmentation, the multiple LN layers and the branch selection strategy based on cosine similarity.
- Unet*: The standard Unet combined with the source domain data augmentation, the multiple LN layers and the branch selection strategy based on cosine similarity.
- MLN-net-E+: Compared to MLN-net-E, a correctional mean parameter is added to the matrices Q_d and Q_t.
- MLN-net+: Compared to MLN-net, a correctional mean parameter is added to the matrices Q_d and Q_t.

Table 4
Segmentation performance comparison with different configurations. DSC and HD are used to evaluate the performance of these methods (best result is in bold for each column). Target DBT (Source FFDM): the models are trained on the domain FFDM and tested on the domain DBT, and Target FFDM (Source DBT): the models are trained on the domain DBT and tested on the domain FFDM.

Datesets: FFDM-DBT	Target DBT (Source FFDM)		Target FFDM (Source DBT)	
Method	DSC (%)	HD (mm)	DSC (%)	HD (mm)
Basenet	46.97	29.17	57.07	26.21
Basenet-S	67.48	25.06	70.93	24.33
MLN-net-R	68.12	25.95	73.03	24.39
Unet-R	53.62	30.10	53.03	31.75
MLN-net-E	78.76	24.02	84.96	22.62
Unet-E	66.18	25.44	66.12	25.39
MLN-net	77.91	23.32	85.52	20.49
Unet*	68.99	26.47	69.07	24.76
MLN-net-E+	78.53	**22.42**	86.12	20.69
MLN-net+	**79.12**	22.75	**89.37**	**19.66**

5.1.1. Efficacy of the source domain data augmentation

The source-similar data augmentation, based on Bézier curves, requires manually initialization of hyperparameters. To evaluate the impact of these hyperparameters, an ablation study was conducted, in which control point pairs for the Bézier curve were discussed in detail within the context of MLN-net. And the parameters of control point that will be chosen are often set to mean or random values of some variables. The Random strategy takes values arbitrarily chosen between 0 and 0.5 for a, while the Mean strategy generates a uniformly, based on the number of control points (i.e., $a_n = 0.5 \cdot n/N$, where n denotes the label of the control

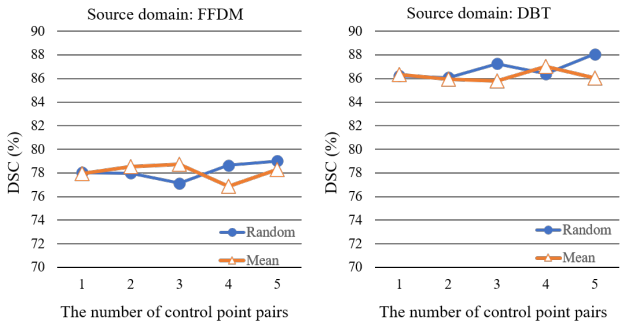

Figure 8: The segmentation performance of MLN-net on the FFDM-DBT dataset based on different numbers of control point pairs for the Bézier curve and different selection strategies of parameters. The vertical axis represents the DSC and the horizontal axis represents different number of control point pairs of the Bézier curve.

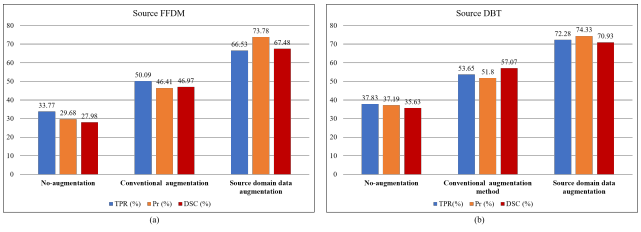

Figure 9: Evaluation comparison of source domain image augmentation method.The experimental results shown in (a) and (b) are from the source domains of FFDM and DBT, respectively.

point pair). Once a is established, a pair of control points can be identified as (a, 1-a) and (1-a, a). Fig. 8 represents the results evaluated for DSC, suggesting that the performance of MLN-net is less impacted by both the number of Bézier curve's control point pairs and the selection strategy of parameters, as long as they are maintained within a sensible range. However, the increase in the number of control point pairs correspondingly expands the training data volume, resulting in a challenging computational problem. And the Random strategy may introduce uncertainty into MLN-net. As a result, we chose two control points and set a at 0.30 and 0.70 Zhou et al. (2022, 2020). And both the start and end points are initialized to (0,0) and (1,1), in accordance with the image's pixel value range.

Based on these parameters settings, we further investigated the effect of the source domain data augmentation on clustered microcalcifications segmentation. The original Swin-Unet was selected as the base network for this experiment to ensure that the results remained untainted by multiple LN layers and varying branch selection strategy. Our control group comprises both a non-augmentation group and a conventional augmentation groups, where conventional augmentation methods consists of random cropping, flipping, and scaling geometric transformations. Domain shifts in mammography primarily manifest as disparities in image quality and appearance, with brightness being a key aspect. Among various modalities, the primary distinguishing factor among the datasets lies in the variation of pixel mapping strategies, leading to pixel brightness differences. Consequently, we considered the source domain data augmentation, employing the monotonic non-linear and grayscale-inversion transformation for medical images, as the absolute intensity values convey information of the underlying substructure Forbes (2012); Buzug (2011). Our proposed augmentation generates data that closely resembles real-world scenarios and enriches the distribution of samples, effectively mitigating the impact of domain shifts. Fig. 9 presents TPR, PR and DSC for the segmentation of testing sets on unseen domains. On the source domain DBT, Swin-Unet, with the proposed source domain data augmentation, achieves a DSC of 70.93%, surpassing the conventional augmentation by 13.86%. Table 4 demonstrates that Basenet-S (augmented with source domain data augmentation) attains a HD of 25.06mm on the target domain DBT. When compared to Basenet, HD decreased by 4.11mm, providing evidence that the proposed source domain data augmentation enhances the model's ability to differentiate ambiguous lesion areas. Being a distance-based metric, the improvement in HD serves as a demonstration of the advantages brought about by the proposed source domain data augmentation in distinguishing ambiguous lesion areas. The visualization results of Fig. 5 offer a more intuitive representation of this advantage.

5.1.2. *Efficacy of the segmentation network with multiple LN layers and the branch selection strategy*

MLN-net employs the segmentation network with multiple LN layers to capture image features and domain information. It also leverages the branch selection strategy to determine the optimal segmentation outcomes. This section scrutinizes the benefits of using these modules by comparing the resulting performance, as shown in Table 4. MLN-net-R and Unet-R models do not include the branch selection strategy. The integration over multi-domain outputs can be performed by summing over the pixel values of segmentation results from different branches. The pixel values greater than or equal to 1 are considered as lesion areas. It can be observed that MLN-net-R (with the source domain data augmentation and multiple LN layers) attains a DSC of 68.12% and a HD of 25.95mm. In comparison to Basenet, it demonstrates a significant enhancement of 21.15% in DSC and a notable 3.22 mm improvement in HD. However, it does not yield a significant improvement compared to Basenet-S (only with domain data augmentation). The results suggest that the primary source of these performance enhancements is attributed to the source domain data augmentation. And the application of multiple normalization layers alone is not sufficient to capture multi-domain distribution information. By the introduction of branch selection strategy, MLN-net demonstrates a noticeable performance improvement in comparison to MLN-net-R. This finding suggests that the

branch selection strategy within MLN-net is closely intertwined with the multiple normalization layers. The multiple LN layers are designed to segregate and record domain-specific distribution information, whereas the branch selection strategy is tailored for establishing the normalization branch based on similarity calculations between the target and source domains. The effective synergy between these components stands as a crucial factor contributing to MLN-net's exceptional performance.

Furthermore, we explored the potential advantages of using cosine similarity in the branch selection strategy. As shown in Table 4, when Q_d and Q_t only encompass mean and variance of data, the cosine similarity does not exhibit significant superiority over Euclidean distance. The introduction of a correctional mean parameter in Qd and Qt provides MLN-net+ an improved segmentation performance on the source domain DBT, with DSC and HD reaching 89.37% and 19.66mm, respectively. This performance signifies an improvement of 4.41% and 2.96mm compared with MLN-net-E. Importantly, this correctional mean parameter is taken from the mean of the input pixels, excluding the minimum pixel value that corresponds to the background region in mammography. In practice, the size of the background region fluctuates significantly with respect to various breast sizes, affecting the global pixel mean. The comparison between MLN-NET+ and MLN-net indicates that integrating additional domain distribution information can lead to an enhancement in performance. Moreover, compared to MLN-Net-E+, MLN-net+ improves DSC and HD from 86.12% and 20.69mm to 89.37% and 19.66mm on the source domain DBT, indicating the benefits of branch selection strategy with cosine similarity in processing high-dimensional data.

In the ablation studies, we also established the baseline models based on Unet framework. We discovered that models based on the Swin-Unet framework outperform those based on Unet framework with a significant gap. The results verify the superior efficacy of MLN-net's feature extraction network, which is rooted in self-attention mechanisms.

5.2. Analysis on params of MLN-net' segmentation network

MLN-net utilizes the Swin-Unet architecture to extract data feature and applies the multiple LN layers to differentiate the characteristics of various domains. These LN layers serve as containers, separating domain distribution information using mean and standard deviation. This distribution information serves as classification standard to determine the possible domain affiliations of the target domain. However, the use of multiple LN layers increases the parameters of the segmentation network. To assess the impact of these multiple LN layers on the network's parameter count, we compared the total number of parameters between Swin-Unet and MLN-net's segmentation network. Both networks take single-channel, 512×512 resolution images as input to eliminate the influence of different inputs. The experiment reveals that Swin-Unet and MLN-net' segmentation network have 83.6M and 84.9M parameters, respectively. MLN-net' segmentation network has 1.3M additional parameters compared to Swin-Unet. Nevertheless, Swin-Unet fails to effectively capture multi-domain data features and address domain shift issues. This necessitates network retraining, which in turn doubles the required parameter count. MLN-net's segmentation network accomplishes the separation of domain distribution information by employing independent multiple LN layers. Obviously the proposed straightforward yet efficient framework has introduced additional layers, but it is worth the extra parameters for the improvement on feature extraction from multi-domain data.

6. Conclusion

In this study, we developed a multi-source medical image segmentation method, MLN-net, for the segmentation of clustered microcalcifications, a key indicator of breast cancer. Combining the source domain data augmentation based on Bézier curves and grayscale-inversion, it improves data diversity dramatically. And a novel segmentation network is constructed using multiple LN layers, so as to extract features from both source-similar and source-dissimilar data. Furthermore, the branch selection strategy is incorporated into MLN-net by utilizing cosine similarity distance to get the best results for target domain. To our knowledge, MLN-net is the first cross-center generalization method for segmenting clustered microcalcifications in breast cancer diagnosis.

We aimed to develop a clustered microcalcifications segmentation method with cross-center generalization ability, which overcomes the binding characteristics inherent to deep learning models and data sources, thus bridging the data gap between imaging methods of different hospitals. MLN-net, no doubt, provides an innovative solution for intelligent breast cancer detection and treatment, with the potential to be applied widely in a clinically meaningful setting.

While the MLN-net has some important practical properties, it suffers from several limitations particularly in the interpretability. Deep learning-based methods are often difficult to understand the rationale behind the results, posing significant challenges for medical diagnoses based on artificial intelligence and impeding the clinical application of these approaches. Consequently, future research should concentrate on improving the interpretability of MLN-net, introducing the key concepts needed for an understanding of the complex models, and assisting healthcare professionals in disease diagnosis and treatment.

7. Acknowledgments

This work is supported by the National Science and Technology Program of China [G2022016011L], the "Ling Yan" Research and Development Project of Science and Technology Department of Zhejiang Province of China [No.2022C03122], the Zhejiang Provincial Natural Science Foundation of China [No.LQ23F030002], the Public Welfare Technology Application and Research Projects of Science and Technology Department of Zhejiang Province of

China [No.LGF21F010004], and Zhejiang Shuren University Basic Scientific Research Special Funds [2023XZ001].

References

Abuduweili, A., Li, X., Shi, H., Xu, C.Z., Dou, D., 2021. Adaptive consistency regularization for semi-supervised transfer learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6923–6932.

Ba, J.L., Kiros, J.R., Hinton, G.E., 2016. Layer normalization. arXiv preprint arXiv:1607.06450 .

Balaji, Y., Sankaranarayanan, S., Chellappa, R., 2018. Metareg: Towards domain generalization using meta-regularization. Advances in neural information processing systems 31.

Bekker, A.J., Shalhon, M., Greenspan, H., Goldberger, J., 2015. Multi-view probabilistic classification of breast microcalcifications. IEEE Transactions on medical imaging 35, 645–653.

Bougourzi, F., Distante, C., Dornaika, F., Taleb-Ahmed, A., 2023. Pdatt-unet: Pyramid dual-decoder attention unet for covid-19 infection segmentation from ct-scans. Medical Image Analysis , 102797.

Butoi, V.I., Ortiz, J.J.G., Ma, T., Sabuncu, M.R., Guttag, J., Dalca, A.V., 2023. Universeg: Universal medical image segmentation. arXiv preprint arXiv:2304.06131 .

Buzug, T.M., 2011. Computed tomography, in: Springer handbook of medical technology. Springer, pp. 311–342.

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., Wang, M., 2023. Swin-unet: Unet-like pure transformer for medical image segmentation, in: Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part III, Springer. pp. 205–218.

Carneiro, G., Nascimento, J., Bradley, A.P., 2017. Automated analysis of unregistered multi-view mammograms with deep learning. IEEE transactions on medical imaging 36, 2355–2365.

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A.L., Zhou, Y., 2021. Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 .

Chong, A., Weinstein, S.P., McDonald, E.S., Conant, E.F., 2019. Digital breast tomosynthesis: concepts and clinical practice. Radiology 292, 1.

Conant, E.F., Zuckerman, S.P., McDonald, E.S., Weinstein, S.P., Korhonen, K.E., Birnbaum, J.A., Tobey, J.D., Schnall, M.D., Hubbard, R.A., 2020. Five consecutive years of screening with digital breast tomosynthesis: outcomes by screening year and round. Radiology 295, 285–293.

Doersch, C., Gupta, A., Efros, A.A., 2015. Unsupervised visual representation learning by context prediction, in: Proceedings of the IEEE international conference on computer vision, pp. 1422–1430.

Dong, J., Cong, Y., Sun, G., Liu, Y., Xu, X., 2020. Cscl: Critical semantic-consistent learning for unsupervised domain adaptation, in: European Conference on Computer Vision, Springer. pp. 745–762.

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

Du, L., Tan, J., Yang, H., Feng, J., Xue, X., Zheng, Q., Ye, X., Zhang, X., 2019. Ssf-dan: Separated semantic feature based domain adaptation network for semantic segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 982–991.

Fan, X., Wang, Q., Ke, J., Yang, F., Gong, B., Zhou, M., 2021. Adversarially adaptive normalization for single domain generalization, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8208–8217.

Forbes, G.B., 2012. Human body composition: growth, aging, nutrition, and activity. Springer Science & Business Media.

Fu, H., Cheng, J., Xu, Y., Wong, D.W.K., Liu, J., Cao, X., 2018. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. IEEE transactions on medical imaging 37, 1597–1605.

Giess, C.S., Pourjabbar, S., Ip, I.K., Lacson, R., Alper, E., Khorasani, R., 2017. Comparing diagnostic performance of digital breast tomosynthesis and full-field digital mammography in a hybrid screening environment. American Journal of Roentgenology 209, 929–934.

Gu, R., Wang, G., Song, T., Huang, R., Aertsen, M., Deprest, J., Ourselin, S., Vercauteren, T., Zhang, S., 2020. Ca-net: Comprehensive attention convolutional neural networks for explainable medical image segmentation. IEEE transactions on medical imaging 40, 699–711.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

He, X., Ying, G., Zhang, J., Chu, X., 2022. Evolutionary multi-objective architecture search framework: Application to covid-19 3d ct classification, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 560–570.

He, Y., Carass, A., Zuo, L., Dewey, B.E., Prince, J.L., 2021. Autoencoder based self-supervised test-time adaptation for medical image analysis. Medical image analysis 72, 102136.

Horvat, J.V., Keating, D.M., Rodrigues-Duarte, H., Morris, E.A., Mango, V.L., 2019. Calcifications at digital breast tomosynthesis: imaging features and biopsy techniques. Radiographics 39, 307–318.

Hossain, M.S., 2022. Microc alcification segmentation using modified u-net segmentation network from mammogram images. Journal of King Saud University-Computer and Information Sciences 34, 86–94.

Ibtehaz, N., Rahman, M.S., 2020. Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation. Neural networks 121, 74–87.

In, T., 2019. Facts & figures 2019: Us cancer death rate has dropped 27% in 25 years. American Cancer .

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International conference on machine learning, PMLR. pp. 448–456.

Jakubovitz, D., Rodrigues, M.R., Giryes, R., 2019. Lautum regularization for semi-supervised transfer learning, in: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, pp. 0–0.

Jiang, P., Huang, S., Fu, Z., Sun, Z., Lakowski, T.M., Hu, P., 2020. Deep graph embedding for prioritizing synergistic anticancer drug combinations. Computational and structural biotechnology journal 18, 427–438.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al., 2023. Segment anything. arXiv preprint arXiv:2304.02643 .

Kowald, A., Barrantes, I., Möller, S., Palmer, D., Murua Escobar, H., Schwerk, A., Fuellen, G., 2022. Transfer learning of clinical outcomes from preclinical molecular data, principles and perspectives. Briefings in Bioinformatics 23, bbac133.

Lee, R.S., Gimenez, F., Hoogi, A., Miyake, K.K., Gorovoy, M., Rubin, D.L., 2017. A curated mammography data set for use in computer-aided detection and diagnosis research. Scientific data 4, 1–9.

Li, Z., Zhang, C., Zhang, Y., Wang, X., Ma, X., Zhang, H., Wu, S., 2022. Can: Context-assisted full attention network for brain tissue segmentation. Medical Image Analysis , 102710.

Liu, J., Zarshenas, A., Qadir, A., Wei, Z., Yang, L., Fajardo, L., Suzuki, K., 2018. Radiation dose reduction in digital breast tomosynthesis (dbt) by means of deep-learning-based supervised image processing, in: Medical imaging 2018: Image processing, SPIE. pp. 89–97.

Liu, Q., Chen, C., Qin, J., Dou, Q., Heng, P.A., 2021. Feddg: Federated domain generalization on medical image segmentation via episodic learning in continuous frequency space, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1013–1023.

Liu, Q., Dou, Q., Heng, P.A., 2020a. Shape-aware meta-learning for generalizing prostate mri segmentation to unseen domains, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23, Springer. pp. 475–485.

Liu, Z., Yang, X., Gao, R., Liu, S., Dou, H., He, S., Huang, Y., Huang, Y., Luo, H., Zhang, Y., et al., 2020b. Remove appearance shift for ultrasound image segmentation via fast and universal style transfer, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1824–1828.

Ma, C., Ji, Z., Gao, M., 2019. Neural style transfer improves 3d cardiovascular mr image segmentation on inconsistent data, in: Medical

Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22, Springer. pp. 128–136.

Miao, J., Wang, X., Wu, Y., Li, W., Zhang, X., Wei, Y., Yang, Y., 2022. Large-scale video panoptic segmentation in the wild: A benchmark, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 21033–21043.

Miao, J., Wei, Y., Wang, X., Yang, Y., 2023. Temporal pixel-level semantic understanding through the vspw dataset. IEEE Transactions on Pattern Analysis and Machine Intelligence .

Miao, J., Wei, Y., Wu, Y., Liang, C., Li, G., Yang, Y., 2021. Vspw: A large-scale dataset for video scene parsing in the wild, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4133–4143.

Mukama, T., Kharazmi, E., Xu, X., Sundquist, K., Sundquist, J., Brenner, H., Fallah, M., 2020. Risk-adapted starting age of screening for relatives of patients with breast cancer. JAMA oncology 6, 68–74.

Pan, W., Song, K., Zhang, Y., Yang, C., Zhang, Y., Ji, F., Zhang, J., Shi, J., Wang, K., 2022. The molecular subtypes of triple negative breast cancer were defined and a ligand-receptor pair score model was constructed by comprehensive analysis of ligand-receptor pairs. Frontiers in Immunology 13.

Perkonigg, M., Hofmanninger, J., Herold, C.J., Brink, J.A., Pianykh, O., Prosch, H., Langs, G., 2021. Dynamic memory to alleviate catastrophic forgetting in continual learning with medical imaging. Nature communications 12, 5678.

Richman, I.B., Hoag, J.R., Xu, X., Forman, H.P., Hooley, R., Busch, S.H., Gross, C.P., 2019. Adoption of digital breast tomosynthesis in clinical practice. JAMA internal medicine 179, 1292–1295.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer. pp. 234–241.

Samala, R.K., Chan, H.P., Hadjiiski, L.M., Cha, K., Helvie, M.A., 2016. Deep-learning convolution neural network for computer-aided detection of microcalcifications in digital breast tomosynthesis, in: Medical Imaging 2016: Computer-Aided Diagnosis, SPIE. pp. 234–240.

Segu, M., Tonioni, A., Tombari, F., 2023. Batch normalization embeddings for deep domain generalization. Pattern Recognition 135, 109115.

Seo, S., Suh, Y., Kim, D., Kim, G., Han, J., Han, B., 2020. Learning to optimize domain specific normalization for domain generalization, in: European Conference on Computer Vision, Springer. pp. 68–83.

Sourati, J., Gholipour, A., Dy, J.G., Tomas-Fernandez, X., Kurugol, S., Warfield, S.K., 2019. Intelligent labeling based on fisher information for medical image segmentation using deep learning. IEEE transactions on medical imaging 38, 2642–2653.

Strehl, A., Ghosh, J., Mooney, R., 2000. Impact of similarity measures on web-page clustering, in: Workshop on artificial intelligence for web search (AAAI 2000), p. 64.

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F., 2021. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians 71, 209–249.

Tarver, T., 2012. Cancer facts & figures 2012. american cancer society (acs) atlanta, ga: American cancer society, 2012. 66 p., pdf. available from.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Advances in neural information processing systems 30.

Wang, J., Yang, Y., 2018. A context-sensitive deep learning approach for microcalcification detection in mammograms. Pattern recognition 78, 12–22.

Wang, S., Yu, L., Li, K., Yang, X., Fu, C.W., Heng, P.A., 2020. Dofe: Domain-oriented feature embedding for generalizable fundus image segmentation on unseen datasets. IEEE Transactions on Medical Imaging 39, 4237–4248.

Wang, S.H., Nayak, D.R., Guttery, D.S., Zhang, X., Zhang, Y.D., 2021. Covid-19 classification by ccshnet with deep fusion using transfer learning and discriminant correlation analysis. Information Fusion 68, 131–148.

Wei, W., Meng, D., Zhao, Q., Xu, Z., Wu, Y., 2019. Semi-supervised transfer learning for image rain removal, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 3877–3886.

Wichakam, I., Chayakulkheeree, J., Vateekul, P., 2018. Deep multi-label 3d convnet for breast cancer diagnosis in dbt with inversion augmentation, in: Tenth international conference on digital image processing (ICDIP 2018), SPIE. pp. 1565–1577.

Zhang, B., Wang, Y., Hou, W., Wu, H., Wang, J., Okumura, M., Shinozaki, T., 2021. Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling. Advances in Neural Information Processing Systems 34, 18408–18419.

Zhang, K., Karanth, S., Patel, B., Murphy, R., Jiang, X., 2022. A multi-task gaussian process self-attention neural network for real-time prediction of the need for mechanical ventilators in covid-19 patients. Journal of Biomedical Informatics 130, 104079.

Zhang, L., Wang, X., Yang, D., Sanford, T., Harmon, S., Turkbey, B., Wood, B.J., Roth, H., Myronenko, A., Xu, D., et al., 2020. Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. IEEE transactions on medical imaging 39, 2531–2540.

Zheng, J., Sun, H., Wu, S., Jiang, K., Peng, Y., Yang, X., Zhang, F., Li, M., 2020. 3d context-aware convolutional neural network for false positive reduction in clustered microcalcifications detection. IEEE Journal of Biomedical and Health Informatics 25, 764–773.

Zhou, K., Yang, Y., Hospedales, T., Xiang, T., 2020. Deep domain-adversarial image generation for domain generalisation, in: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 13025–13032.

Zhou, Z., Qi, L., Yang, X., Ni, D., Shi, Y., 2022. Generalizable cross-modality medical image segmentation via style augmentation and dual normalization, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 20856–20865.

Zuckerman, S.P., Sprague, B.L., Weaver, D.L., Herschorn, S.D., Conant, E.F., 2020. Multicenter evaluation of breast cancer screening with digital breast tomosynthesis in combination with synthetic versus digital mammography. Radiology 297, 545–553.