

Self-Supervised Hyperbolic Metric Few-Shot Learning for Occult Lymph Node Metastasis Prediction in NSCLC

Haizhou Xu, Jiaqi Wu, Yujia Yu, Wenkai Huang, *Member, IEEE* and Jiong Ni

Abstract—Predicting occult lymph node metastasis in non-small cell lung cancer (NSCLC) patients before surgery is crucial for determining appropriate treatment plans. However, predicted occult lymph node metastases in NSCLC patients have been under-studied due to various challenges, such as the complexity and variability of lymph node characteristics. In addition, there is no publicly available pathologically confirmed predictive dataset of occult lymph node metastases in NSCLC patients for training and testing predictive algorithms. To address this challenge we retrospectively collect occult lymph node metastasis dataset (TJ-OLNM) that meet the criteria from NSCLC patients who undergo chest CT scans at Tongji Hospital of Tongji University between 2016 and 2021. Additionally, we introduce a novel self-supervised hyperbolic metric few-shot learning method, named Occult Lymph Node Metastasis Network (OLNM-Net), for accurate prediction of occult lymph node metastasis. In OLN-Net, a self-supervised pretrained module is first used to learn preliminary lymph node imaging features, and then a hyperbolic theory-based few-shot metric learning module is used to predict occult lymph node metastasis in NSCLC. We performed a comprehensive evaluation of the state-of-the-art prediction model and our OLN-Net on the constructed TJ-OLNM dataset. The experimental results show that our OLN-Net yields better performance in predicting occult lymph node metastasis in TJ-OLNM than traditional and other deep learning methods. This explores the possibility of preoperative prediction of occult lymph node metastasis in NSCLC from CT imaging features, providing a new scheme for the study of NSCLC disease.

Index Terms—Lymph nodes, Hyperbolic space, Self-supervised learning, Metric learning, Few-shot learning.

I. INTRODUCTION

LUNG cancer is among the most prevalent and deadly diseases worldwide [1], with non-small cell lung cancer (NSCLC) being the most common type [2]. NSCLC is characterized by its rapid progression, poor prognosis, and tendency to develop chemotherapy resistance. Lymph node metastasis is the most common route of metastasis in NSCLC.

This work was supported by the Guangzhou Youth Science and Technology Education Project under Grant KP2023243. (Corresponding authors: Wenkai Huang; Jiong Ni.)

Haizhou Xu, Yujia Yu, and Wenkai Huang are with the School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China (e-mail: 2007700059@e.gzhu.edu.cn; 2007700026@e.gzhu.edu.cn; smallkat@gzhu.edu.cn)

Jiaqi Wu and Jiong Ni are with the Department of Radiology, Tongji Hospital, School of Medicine, Tongji University, Shanghai, 200065, China. (e-mail: wujiaqitj@sina.com; ni.jiong@163.com).

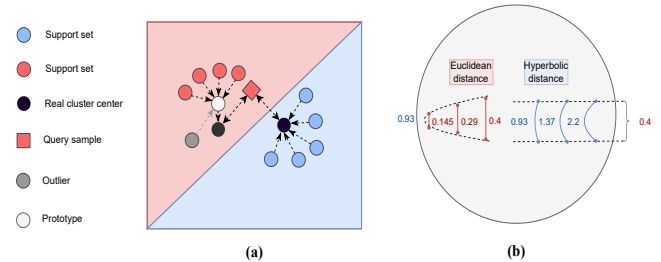


Fig. 1. (a) For the prototype embedding in Euclidean space, the query sample of the red class is misclassified as blue by the classifier because the outlier pulls the prototype away from the real clustering center in the embedding space. (b) Comparison between Euclidean distance and hyperbolic distance. In hyperbolic space (2-dimensional Poincaré disk), due to the difference in degree gauge, the farther the distance from the center of the circle, the greater the spatial curvature and the greater the geodesic or distance between points, which can better represent the potential hierarchical clustering relationship compared with Euclidean space.

On chest CT scans, lymph nodes with a short axis of less than 10 mm and confirmed as metastatic after surgery are defined as occult lymph node metastasis [3]. Due to the presence of occult lymph node metastasis, conventional examination methods have limited value in determining lymph node staging. Predicting the occurrence of occult lymph node metastasis before surgery is of paramount clinical significance in accurately assessing regional lymph node metastasis and assisting clinicians in selecting the most appropriate treatment plan for NSCLC patients.

Currently, preoperative assessment of lymph node metastasis in NSCLC patients includes CT, positron emission tomography/X-ray computed tomography (PET/CT), ultrasound-guided fine needle aspiration biopsy, and mediastinoscopy [4]–[7]. CT is the most widely used detection method in clinical practice. However, conventional CT mainly evaluates the lymph node metastasis status through morphological indicators, usually using a short-axis diameter of 10 mm as the evaluation indicator, resulting in a single quantitative indicator with limited sensitivity and specificity. PET/CT can better evaluate the metastasis status of lymph nodes. However, NSCLC patients often have complications, which can readily result in false-negative results [8], [9], and the image acquisition scanning is time-consuming with a high radiation dose and high cost, making it difficult for

patients [10]. Ultrasound-guided fine needle aspiration biopsy and mediastinoscopy, as invasive procedures, can cause various complications during surgery that can have a negative impact on patients without lymph node metastasis [11]. In short, current methods for preoperative lymph node diagnosis are limited, and there exist lymph nodes with postoperative pathologically confirmed metastases but no suspicious hints and short diameter (less than 10 mm) on preoperative CT or PET/CT imaging, i.e., occult lymph nodes. It is necessary to find a reliable imaging method to predict occult lymph node metastasis in NSCLC patients.

In recent years, thanks to the rapid development of artificial intelligence, medical research fields have introduced machine learning methods. These methods can extract many data representations from imaging and clinical data, providing a high-dimensional and exploratory feature space. Zhong et al. [12] used imaging features, clinical data, tumor pathological subtype, diameter, and location to construct a multi-parameter logistic regression model to predict mediastinal lymph node metastasis in lung adenocarcinoma. Cong et al. [13] used minimum absolute shrinkage and selection operator (LASSO) constrained logistic regression as qualitative CT morphological features and radiomics features to construct a model to predict lymph node metastasis in NSCLC patients. Ferreira Junior et al. [14] used multiple machine learning models to predict the histopathology and metastasis of lung cancer by using radiomics CT features. However, the aforementioned studies have focused solely on the research of lymph node metastasis and have not delved into the possibilities of predicting occult lymph node metastasis. In contrast to conventional lymph node metastasis, occult lymph node metastasis often lacks evident characteristics, making it challenging to observe the metastatic condition through preoperative imaging features. Furthermore, compared to cases of lymph node metastasis, occult lymph node metastasis cases are relatively rare, which presents difficulties in obtaining large amounts of training data and hinders the application of fully supervised deep models.

To address the problem of data scarcity, in recent years, researchers have proposed many methods for few-shot learning (FSL) [15]–[19]. The aim of FSL is to extract more potentially beneficial information from limited data for model training, thereby obtaining a better model. Most currently proposed FSL methods rely on a large amount of annotated data from the training class to learn image representations that generalize well to the invisible test class. However, such a training mechanism is impractical in medical imaging scenarios where annotations are scarce.

To mitigate the scarcity of imaging datasets for predicting occult lymph node metastasis and address the challenge of effective prediction with limited samples, we make the following contributions in this work:

- 1) For the first time in the field of medical image analysis, we built an occult lymph node metastasis prediction CT image dataset, with accurate manual annotation of occult lymph nodes, to promote relevant research in occult lymph node metastasis prediction.
- 2) We propose a novel few-shot occult lymph node metastasis prediction model called OLN-Net. In our model,

we first employ the self-supervised pretrained module named (MCAE) to extract the initial imaging features of lymph nodes. Then, we utilize the Hyperbolic Prototype Embedding Network (HPENet) module to further extract imaging features of occult lymph nodes, reducing the gap between the training data and the target data, aiming to achieve more accurate prediction results.

- 3) We give a full evaluation/benchmarking of occult lymph node metastasis prediction. Experiments show that the proposed OLN-Net works robustly and yields accurate prediction.

II. RELATED WORKS

A. Few-Shot Learning

At present, existing methods for addressing few-shot image classification tasks can be divided into two categories: optimization-based and metric-based. The optimization-based approach aims to quickly adapt to new tasks by learning effective initial model parameters. Finn et al. [15] proposed a model-agnostic meta-learning (MAML) algorithm that trains on different meta-tasks via multi-step gradient updates. Its goal is to improve the basic network model initialization state and quickly achieve good generalization performance on new tasks. Hu et al. [16] proposed a two-stage MAML algorithm compatible with noise mechanisms, alleviating overfitting during training.

The metric-based approach learns the similarity between samples by mapping images to a specific feature space and classifying them based on the relevant distance metrics in the space, with the choice of distance metric being particularly crucial. Snell et al. [17] proposed the prototypical network, which learns an embedding space by computing the average of the support set to obtain a prototype representation for each class and uses Euclidean distance to measure the similarity between the test image and the prototype. Vinyals et al. [18] proposed a matching network that compares the similarity between support set embeddings and test sample embeddings by using cosine distance. Liu et al. [19] combined the prototype network and relation network by classifying new categories based on the minimum distance between the query image and each class's prototype.

In the field of medical image analysis, FSL methods have been relatively underexplored. A. K. Mondal et al. [20] proposed a few-shot 3D multimodal medical segmentation model based on generative networks that uses unannotated data to prevent overfitting through generative adversarial networks. Chen et al. [21] proposed a generative network-based one-shot MRI segmentation method that transfers images from CT space to MRI space and uses generative adversarial learning to guide the training of the MRI-based segmentation network. Paul et al. [22] proposed a set-based learning method for chest X-ray diagnosis. Other applications of FSL have been proposed for brain imaging pattern recognition [23], voxel-wise medical image segmentation [24], and medical image classification [25].

B. Self-Supervised Learning

Self-supervised learning has received considerable attention in computer vision. It is used to transform unsupervised learning problems into supervised learning problems in different ways to obtain a pretrained model for further supervised training. Currently, most self-supervised learning employs one of two approaches: pretext task-based methods and contrastive learning-based methods. Pretext task-based methods typically learn image representations through manually designed tasks such as image colorization [26], image inpainting [27], or geometric transformation recognition [28]. Contrastive learning-based methods [29]–[31] usually learn representations of images by measuring the similarity between two or more views.

The learning method in self-supervised masked reconstruction is gradually gaining in popularity. He et al. [32] were the first to apply the masked model to the field of vision. They designed an asymmetric encoder-decoder structure and learned image representation by randomly masking patches of input images and reconstructing missing pixels, achieving outstanding self-supervised pretraining effects. Zhou et al. [33] applied self-supervised masked reconstruction to medical image analysis tasks and achieved excellent performance in these tasks by mining the inherent features of images through data with the same target task.

Although self-supervised learning models have achieved remarkable results, almost all current self-supervised models require a large amount of data for pretraining, and the size of the model parameters is extremely large. It is difficult to train a model for special medical cases with very little data. Therefore, reducing the size of the pretrained model and effectively mining potential features, as well as designing reasonable finite strategies to reduce the gap between the model and downstream tasks, are still major challenges.

C. Hyperbolic Embeddings

The inherent embedding hierarchy in hyperbolic space, such as tree structures, endows it with low distortion capabilities [34], [35]. This has made hyperbolic embeddings increasingly popular in the field of natural language processing. Nickel et al. [36] were the first to propose using the Poincaré model to learn embeddings, demonstrating that Poincaré embeddings can learn the latent hierarchical structure in data and that they are superior to Euclidean embeddings in terms of representational and generalization capabilities. A large body of work in natural language processing has employed hyperbolic embeddings to tackle various tasks, including text classification [37], text generation [38], word embeddings [39], [40], and entity typing [41], [42].

However, due to the different nature of data in natural language processing versus computer vision, it is not possible to directly extend hyperbolic space to the visual domain. Khrulkov et al. [43] observed the potential hierarchical relationships between images and utilized the second last layer of a deep neural network to learn higher-level semantic representations for image embedding. Then, the image classification task can be viewed as a certain type of geometry in the embedding space, where pairs of classes are separated by

Euclidean hyperplanes. Grattarola et al. [44] constructed a hyperbolic adversarial autoencoder that outperforms autoencoders based on Euclidean and non-Euclidean geometries in semi-supervised classification, link prediction, and molecule generation tasks. Peng et al. [45] introduced Poincaré ST-GCN for skeleton-based action recognition, modeling input sequences (skeletons) using graph structure, and projecting the structural features extracted by the network into hyperbolic space for action recognition. In these works, one observes the potential of hyperbolic embedding in computer vision. In this paper, we explore the effectiveness of embedding medical image features into hyperbolic space for the first time.

III. METHOD

The central idea of this paper is to effectively extract 3D discriminative features of occult lymph nodes by learning the image representations from numerous unlabeled CT data in a self-supervised learning manner. Then, the finitute strategy of few shot metric learning is used to reduce the gap with the target task. In Fig. 2, the workflow of the proposed HMCAE is presented. In the first step, a lightweight self-supervised mask-based pretraining model called MCAE is created. It consists of a simple asymmetric convolutional encoder-decoder. Similar to MAE [32], CT image data similar to the target task is divided into non-overlapping 3D patches. However, unlike MAE's random sampling mask strategy, there may be a large amount of irrelevant background information in medical images, which may further affect the reconstruction effect. Therefore, similar to sentence completion in natural language processing, patches outside the Region of Interest (ROI) are marked as padding tokens, and when calculating the loss, they are ignored. Then, ROI patches are randomly masked, and all patches are input into MCAE for reconstruction pretraining task. The second step is to fine-tune the target task using a few-shot learning strategy. In this step, we propose the Hyperbolic Prototype Embedding Network (HPENet). First, the MCAE encoder is used to extract image features, and then these features are embedded and mapped to the hyperbolic space. A hyperbolic prototype loss function is constructed to perform metric learning in the hyperbolic space.

A. Self-supervised pretrained

During the self-supervised pretraining step, we developed a lightweight model called MCAE. MCAE uses a self-supervised masked reconstruction training strategy to predict the original image from randomly masked images, thereby obtaining an intuitive representation of the occult lymph node dataset and providing good prior knowledge for the next few-shot classification task. The entire process is divided into the following parts.

1) *Masking*: The masking module is primarily responsible for converting the 3D volume image into a 3D volume mask through random masking. For a 3D image $X \in \mathbb{R}^{C \times H \times W \times D}$, it is first reshaped into a sequence of 3D patches $X_p \in \mathbb{R}^{N \times (C \cdot P^3)}$, where C is the number of channels of the input image, (H, W, D) are the length, width, and height of the input image, P is the size of the non-overlapping patch blocks used

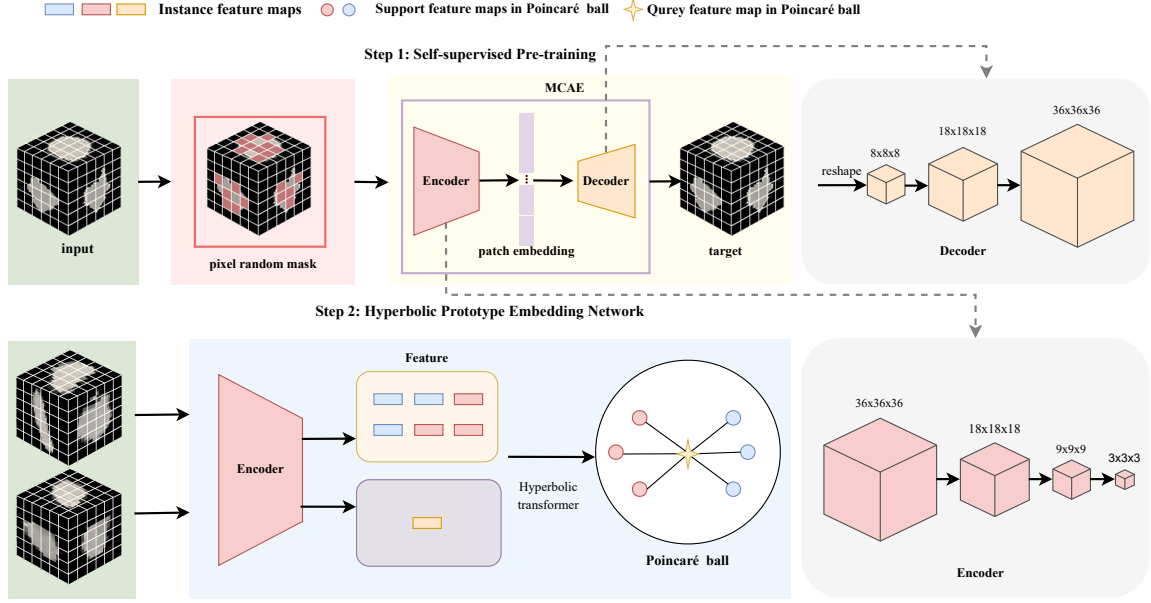


Fig. 2. Architecture of the proposed OLM-Net network for NSCLC occult lymph node metastasis prediction. It consists of two parts, MCAE (Step 1) and HPENet (Step 2). MCAE is designed to initially extract the image features of the lymph nodes. HPENet is then introduced as a fine-tuning network to obtain the final prediction results.

to divide the image, and $N = H \times W \times D/P^3$ is the length of the 3D patch sequence.

Due to the difference in information density distribution between medical images and images in the real world, the global random masking strategy of MAE may introduce unnecessary noise interference in medical images. Therefore, we use a pixel random mask strategy, where patches consisting entirely of background pixels are labeled as background patch P_b and ignored during the masking process. Patches that contain pixels in the ROI region are marked as P_r , and a random masking operation is performed during the masking process, retaining only a small subset (e.g., 25%). Finally, all patches are unpatchified into $X_{input} \in \mathbb{R}^{C \times H \times W \times D}$, where the length of the unmasked patch sequence is $N_{visible} = (1 - \text{maskingratio})N_r$, the length of the masked patch sequence is $N_{mask} = N - N_{visible}$, and the values of the patches marked as background and masked are set to 0.

2) MCAE Encoder: The MCAE encoder is a lightweight three-layer 3D convolutional neural network with a feature-embedding layer, which embeds the lymph node features extracted by the 3D convolutional neural network into the feature space. To maintain the spatial position of each patch, we process all patches (including background patches, visible patches, and masked patches) in the encoder.

3) MCAE Decoder: In the MCAE decoder, we employ an asymmetric structure design, utilizing two deconvolutional layers to reconstruct the input lymph node image from the feature vector embedded in the encoder. This lightweight asymmetric design both shortens the pretraining time and encourages the encoder to explore more advantageous features at a deeper level.

4) Loss Function in MCAE: To train the MCAE, we construct a loss function based on the prediction of masked pixel/voxel values, using mean squared error. Specifically, for the output of the MCAE encoder, P_{out} , we first extract the masked patch using Eqs. (1) and (2):

$$P_{mask} = P \times (1 - \text{mask}) \quad (1)$$

$$\tilde{P}_{mask} = P_{out} \times (1 - \text{mask}) \quad (2)$$

where P represents the patch that was originally divided from the image, and the mask is a binary matrix. The value of the masked patch is set to 0, while the values of other patches (including background patches and visible patches) are set to 1. Then, the mean square error loss is calculated as shown in Eq. (3):

$$\mathcal{L}(P_{mask}, \tilde{P}_{mask}) = \frac{1}{N_{mask}} (P_{mask}^i - \tilde{P}_{mask}^i)^2 \quad (3)$$

where P_{mask}^i and \tilde{P}_{mask}^i denote the patch of the i -th mask and the patch of the reconstructed mask, respectively.

B. Hyperbolic Geometry

Before proceeding further, let us briefly introduce hyperbolic geometry, as this will be beneficial for understanding and derivation in the next subsection.

1) Basics of Riemannian Geometry: A differential manifold of n dimensions is locally homeomorphic to a Euclidean space \mathbb{R}^n , that is, compact Hausdorff topological space \mathcal{M} . For each point, $p \in \mathcal{M}$, an n -dimensional vector space $\mathcal{T}_p\mathcal{M}$ can be defined as the tangent space, which is homeomorphic to \mathbb{R}^n .

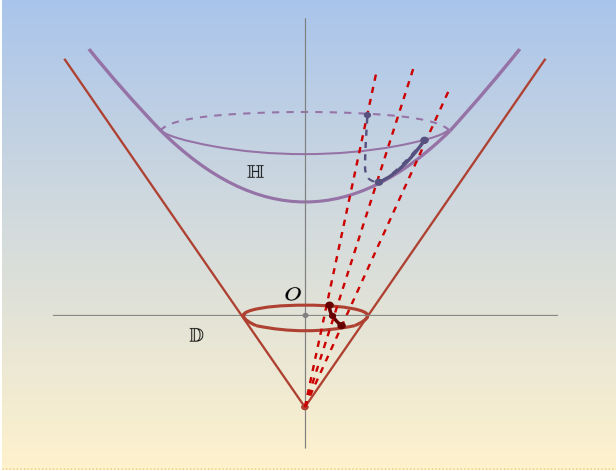


Fig. 3. A bijection between the hyperbolic space model \mathbb{H}^n and the Poincaré model \mathbb{D}^n is induced by projecting a point $p = (0, \dots, 0, -1)$ onto the hyperbolic surface.

The tangent space $\mathcal{T}_p\mathcal{M}$ can be briefly defined as all curves $c : (-\sigma, \sigma) \rightarrow \mathcal{M}$ such that $c(0) = p$ and $\sigma > 0$.

A Riemannian manifold [46] is a differential manifold with a metric tensor, typically represented as (\mathcal{M}, g) , where \mathcal{M} is the manifold and g is the metric tensor. The introduction of the metric tensor restores natural geometric concepts to differential manifolds. Euclidean space can be viewed as a manifold \mathbb{R}^n equipped with the Euclidean metric tensor $g^E = I_n$ at each tangent space $\mathcal{T}_p = \mathbb{R}^n$.

2) Hyperbolic Space: Hyperbolic space is a type of Riemannian manifold with constant negative curvature. It has five isometric models for modeling, namely the hyperboloid (Lorentz) model, the Poincaré ball model, the Poincaré half-space model, the Klein model, and the hemisphere model. In this study, we use the Poincaré ball model.

The Poincaré model is obtained by linearly projecting every point of the hyperbolic surface model \mathbb{H}^n onto the hyperplane $x_0 = 0$ through point $p = (0, \dots, 0, -1)$, as shown in Fig. 3. An n -dimensional Poincaré ball ($\mathbb{D}^n, g^{\mathbb{D}}$) is defined as an n -dimensional Poincaré manifold $\mathbb{D}^n = \{x \in \mathbb{R}^n : \|x\| < 1\}$ equipped with the following Riemannian metric tensor:

$$g_x^{\mathbb{D}} = \lambda_x^2 g^E \quad (4)$$

where λ_x^2 is the conformal factor:

$$\lambda_x^2 = \frac{2}{1 - \|x\|^2} \quad (5)$$

The geodesic distance between two points $u, v \in \mathbb{D}^n$ on the Poincaré ball model can be obtained by the following formula:

$$d_{\mathbb{D}}(u, v) = \cosh^{-1} \left(1 + \frac{2(\|u - v\|^2)}{(1 - \|u\|^2)(1 - \|v\|^2)} \right) \quad (6)$$

C. Hyperbolic Arithmetic Operations

The non-Euclidean nature of hyperbolic space makes it impossible to directly apply basic arithmetic operations such as vector addition and multiplication. One simple method for performing these computations is to use the tangent space for approximate calculations. Another good option is to adopt the Möbius rotation vector space, which provides many standard operations for hyperbolic space. Similar to the previous discussion, we introduce an additional hyperparameter c to adjust the curvature of the Poincaré ball. An n -dimensional Poincaré ball is redefined as $\mathbb{D}_c^n = \{x \in \mathbb{R}^n : c\|x\| < 1, c \geq 0\}$ equipped with a Riemannian metric tensor $g_x^{\mathbb{D}} = \lambda_x^{c^2} g^E$, where the conformal factor $\lambda_x^{c^2}$ is defined as:

$$\lambda_x^{c^2} = \frac{2}{1 - c\|x\|^2} \quad (7)$$

Next, we introduce several fundamental arithmetic operations that are crucial to our work.

1) Möbius addition: Möbius addition is a generalization of addition in Euclidean space to hyperbolic space. In the rotation vector space, the \oplus_c operation between u and v in model \mathbb{D}_c^n is defined as:

$$u \oplus_c v = \frac{(1 + 2c\langle u, v \rangle + c\|v\|^2)x + (1 - c\|u\|^2)v}{1 + 2c\langle u, v \rangle + c^2\|u\|^2\|v\|^2} \quad (8)$$

2) Geodesic distance: The geodesic distance between two points u and v in the model \mathbb{D}_c^n is defined as:

$$d_{\mathbb{D}}^c(u, v) = \frac{2}{\sqrt{c}} \tanh^{-1}(\sqrt{c}\| -u \oplus_c v \|) \quad (9)$$

3) Hyperbolic average: The extension of the arithmetic mean in Euclidean space to hyperbolic space corresponds to the Einstein midpoint, which has the simplest form in the Klein model. Therefore, we first apply the following transformation to map the point $p_{\mathbb{D}}$ on the Poincaré model to the point $p_{\mathbb{K}}$ on the Klein model.

$$p_{\mathbb{K}} = \frac{2p_{\mathbb{D}}}{1 + c\|p_{\mathbb{D}}\|^2} \quad (10)$$

Next, the hyperbolic mean can be obtained by computing the Einstein midpoint [47], which is defined as follows:

$$\mu_{\mathbb{K}} = \frac{\sum_{i=1}^n \gamma_i p_{\mathbb{K}}^i}{\sum_{i=1}^n \gamma_i} \quad (11)$$

where $\gamma_i = \frac{1}{\sqrt{1 - c\|p_{\mathbb{K}}^i\|^2}}$ is the Lorentz factor. Finally, we can easily transform the computed midpoint back to the Poincaré model to obtain the average of the Poincaré model, since the Klein model and Poincaré model are isomorphic. This transformation is defined as follows:

$$\mu_{\mathbb{D}} = \frac{\mu_{\mathbb{K}}}{1 + \sqrt{1 - c\|\mu_{\mathbb{K}}\|^2}} \quad (12)$$

4) Exponential and logarithmic maps: The above are the operational rules defined in hyperbolic space. To perform these operations, we also need to define a bijection between the Euclidean space \mathbb{R}^n and the hyperbolic space \mathbb{D}_c^n . The exponential map $\exp_u^c(v)$ defines a transformation from the Euclidean space to the Poincaré model, defined as follows:

$$\exp_u^c(v) = u \oplus_c \left(\tanh\left(\sqrt{c}\frac{\lambda_x^{c^2}\|v\|}{2}\right) \frac{v}{\sqrt{c}\|v\|} \right) \quad (13)$$

Its inverse transformation, the logarithmic map $\log_u^c(v)$, is defined as follows:

$$\log_u^c(v) = \frac{2}{\sqrt{c}\lambda_x^{c^2}} \tanh^{-1}(\sqrt{c}\| -u \oplus_c v \|) \frac{-u \oplus_c v}{\| -u \oplus_c v \|} \quad (14)$$

D. Hyperbolic Prototype Embedding Network

1) *Overview*: The Hyperbolic Prototype Embedding Network proposed in this paper consists of two main parts: (1) an encoder composed of convolutional neural networks, denoted as $\mathcal{F} : \mathcal{X} \rightarrow \mathbb{R}_c^n$, which extracts features from images and embeds them into an n-dimensional Poincaré ball, and (2) hyperbolic similarity calculation, which computes similarities in hyperbolic space for image classification.

2) *Hyperbolic Prototype Embedding Network*: In FSL, the problem of classifying with a small number of samples is typically represented as an N -way K -shot classification problem. The training set is also known as the support set, and it contains N classes, with each class having K instances. Usually, K is a small number such as 1 or 5.

Given support set $S = \{S_1, \dots, S_N | S_i \in \mathbb{R}^{K \times C \times H \times W \times D}\}$, where C , H , W , and D denote the channel, height, width, and depth of input images, respectively, we first train a convolutional neural network to encode them into an n-dimensional feature space, i.e., $\mathcal{X} = \mathcal{F}_\phi(S)$, where $\mathcal{X} = \{\mathcal{X}_1, \dots, \mathcal{X}_N | \mathcal{X}_i \in \mathbb{D}_c^n\}$, and ϕ represents the learnable parameters of the network. Each hyperbolic prototype s_k is defined as the mean vector of Poincaré ball embeddings of the support points belonging to its category, which can be computed as follows:

$$s_k = \mu_{\mathbb{D}}(\mu_{\mathbb{K}}(p_{\mathbb{K}}(exp_u^c(\mathcal{X}_k))) \quad (15)$$

where $p_{\mathbb{K}}(\cdot)$, $\mu_{\mathbb{D}}(\cdot)$ are operations that transform between the Poincaré model and the Klein model, $exp_u^c(\cdot)$ is a transformation from Euclidean space to the Poincaré model, and $\mu_{\mathbb{K}}(\cdot)$ is the Einstein midpoint operation, which calculates the mean in hyperbolic space.

For a given query set $Q = \{Q_1, \dots, Q_N | Q_i \in \mathbb{R}^{B \times C \times H \times W \times D}\}$, where B is the number of instances per category in the query set, we map them to n-dimensional feature vectors on the Poincaré ball:

$$q_k = exp_u^c(\mathcal{F}_\phi(Q_k)) \quad (16)$$

We can then use Eq. (6) to compute the distance between the query sample q and each support set s of each class. Finally, we can calculate the probability of the current query sample q_i belonging to each hyperbolic prototype:

$$p_\phi(y = k | q_i) = \frac{exp(-d_{\mathbb{D}}(q_i, s_k))}{\sum_{k'} exp(-d_{\mathbb{D}}(q_i, s_{k'}))} \quad (17)$$

Assuming the true class of the sample is k , we can define the loss function as

$$\mathcal{L}(\phi) = -\log p_\phi(y = k | q_i) \quad (18)$$

We aspire to minimize the loss function, which is equivalent to maximizing the logarithm, that is, maximizing the probability of correctly classifying q_i . By minimizing the loss function, we can optimize the Hyperbolic Prototype Embedding Network.

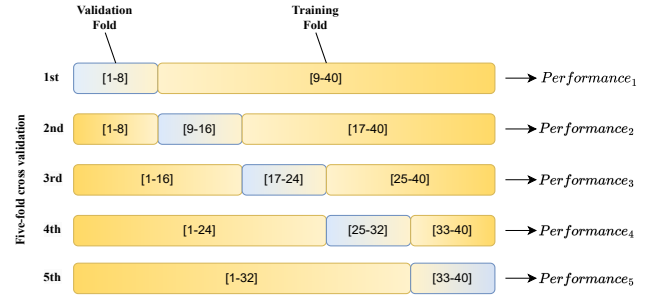


Fig. 4. Five-fold cross validation.

IV. EXPERIMENTS

A. Dataset

This article conducts self-supervised pretraining on publicly available lymph node datasets [48]–[50] and comprehensively evaluates the proposed method on the TJ-OLNM Dataset, a lymph node metastasis dataset collected by our team. Below are detailed descriptions of the two datasets.

1) *LN Dataset*: The LN Dataset comprises CT images of 388 mediastinal lymph nodes from 90 patients and 595 abdominal lymph nodes from 86 patients. The lymph node locations were annotated by radiologists with the National Institutes of Health Clinical Center, and the lymph node segmentation masks were manually traced and segmented based on the methodology described in [7].

2) *TJ-OLNM Dataset*: The TJ-OLNM Dataset was retrospectively collected for this study by selecting eligible cases from chest CT scans of lung cancer patients of Tongji Hospital affiliated with Tongji University. They were collected between 2016 and 2021.

Inclusion criteria for lymph node metastasis: (1) pathologically confirmed primary lung cancer; (2) CT scans of the chest showing non-suspicious target lymph nodes (short-axis diameter <1.0 cm), with postoperative pathological confirmation of tumor metastasis within the lymph nodes; and (3) complete clinical data of patients.

Exclusion criteria for lymph node metastasis: (1) pathologically confirmed primary lung cancer; (2) postoperative pathological confirmation of the absence of tumor metastasis within the target lymph nodes; (3) incomplete clinical patient data.

Our team imported DICOM standard format images of chest CT scans into ITK-SNAP, and a radiologist manually delineated the included lymph nodes layer by layer to generate regions of interest (ROIs). A total of 40 cases met the inclusion criteria.

B. Implementation Details

The method proposed in this study was implemented using the PyTorch framework. All experiments were conducted on a GeForce RTX 3090 GPU with 24 GB memory. As the data were obtained from different devices with varying spacing,

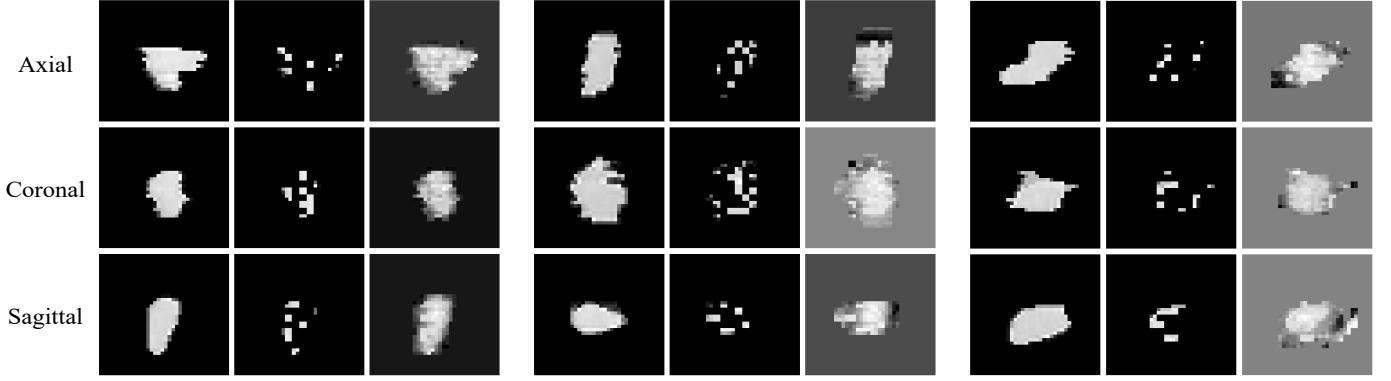


Fig. 5. Example results of self-supervised pre-training. For each triplet, we show the original image (left), the masked image (middle), and the reconstructed image (right). Each row corresponds to one of the three views of the lymph node, with a masking rate of 75%.

we resampled them to a uniform size of 1 mm. Finally, we extracted a $36 \times 36 \times 36$ region at the center of each lymph node based on the manually annotated masks for subsequent experimental procedures.

During the self-supervised pretraining process, we set the initial learning rate as $l = 0.01$ for different masking rates. We then used a fixed-step decay strategy, decaying the learning rate by 0.1 every 100 epochs, until the learning rate remained $l \geq 0.00001$. We set the batch size as $B = 64$ and the maximum iteration epoch as $E = 1000$ and utilized early stopping and 12 regularization methods. Our team used the Adam optimizer to update the network's weights.

During the fine-tuning stage of few-shot learning, in order to comprehensively evaluate the proposed model, we divided the TJ-OLNM dataset equally into five folds for five-fold cross-validation, as shown in Fig. 4. In each fold, the training set was further randomly sampled to obtain support and query sets, and the Hyperbolic Prototype Embedding Network was trained with both 1-shot and 5-shot configurations. The entire test set was used as the query set to evaluate the performance of the model. The hyperparameters of the proposed method and the comparison methods were determined by manually searching for different combinations, and the best-performing models were retained after repeated parameter training and model testing for analysis.

This article employs the mean and standard deviation of six performance metrics to evaluate the obtained prediction results, including accuracy, specificity, sensitivity, precision, F1 score, and area under the ROC curve (AUC). Accuracy represents the proportion of samples correctly classified by the model; specificity represents the ability to correctly identify negative samples; sensitivity represents the ability to correctly identify positive samples; and precision represents the ability to predict correctly among positive samples. The F1 score combines precision and recall: it is the harmonic mean of precision and recall. AUC is a metric that comprehensively considers specificity and sensitivity and thus reflects the expected generalization performance of the model. The formulas for calculating these metrics are as follows:

$$Sensitivity = \frac{TP}{TP + FN} \quad (19)$$

$$Specificity = \frac{TN}{TN + FP} \quad (20)$$

$$Accuracy = \frac{TP + FN}{TP + FP + TN + FN} \quad (21)$$

$$Precision = \frac{TP}{TP + FP} \quad (22)$$

$$F1score = \frac{2 \times Sensitivity \times Precision}{Sensitivity + Precision} \quad (23)$$

where TN and TP represent the number of correctly identified lymph nodes with and without metastasis, respectively. FP and FN represent the number of incorrectly identified lymph nodes with and without metastasis, respectively.

C. Results of Self-supervised Pretraining

Fig. 5 illustrates the reconstruction results of MCAE with a masking rate of 75%. The three rows show the axial, coronal, and sagittal sections of the CT images, respectively. The three columns show the original image (left), the masked image (middle), and the reconstructed image (right). The results show that MCAE can effectively restore patch blocks that have been randomly masked by pixel masking strategies. It should be noted that since the loss function of MCAE only acts on the masked patch blocks, the reconstructed images may appear somewhat blurry. However, the ultimate goal of MCAE is not to generate high-quality reconstruction results but rather to enable the encoder module to learn the imaging features of lymph nodes, which will benefit the subsequent task of classifying occult lymph nodes.

D. Comparisons in OLN Prediction

Table I presents the accuracy, sensitivity/recall, specificity, precision, F1 score, and AUC of the proposed HMCAE model and other methods for predicting occult lymph node metastasis. Each method was evaluated with five independent repetitions, and the average values are reported. Among supervised learning methods, this study evaluated classical CNN models, such as 3D ResNet series (3D ResNet-18, 3D ResNet-34, 3D ResNet-50, 3D ResNet-101), and 3D DenseNet-121 on the TJ-OLNM Dataset. As shown in Table I, with the increase of model depth and complexity, sensitivity, accuracy,

TABLE I

PERFORMANCE OF DIFFERENT METHODS ON TJ-OLNM DATASET FOR OCCULT LYMPH NODE METASTASIS PREDICTION

Inference With Fully Supervised Learning						
Method	SEN (%)	SPE (%)	ACC (%)	AUC (%)	Precision (%)	F1-score (%)
3D ResNet-18	84.00	69.00	76.50	71.50	73.13	78.11
3D ResNet-34	88.00	70.00	79.00	73.30	74.42	80.59
3D ResNet-50	85.00	74.00	79.50	74.60	76.66	80.40
3D ResNet-101	76.00	80.00	78.00	76.25	79.41	77.38
3D DenseNet-121	68.00	85.00	76.50	76.65	82.33	74.14
Inference With Transfer Learning (MedicalNet)						
3D ResNet-18	64.00	86.00	75.00	72.75	82.14	71.89
3D ResNet-34	66.00	85.00	75.50	74.15	81.47	72.91
3D ResNet-50	65.00	85.00	75.00	74.95	81.25	72.22
Inference With K = 1 Shot						
3D ProtoNet	76.00	73.00	74.50	71.95	74.06	74.46
3D MatchingNet	79.00	82.00	80.50	78.40	83.48	79.74
HMCAE	83.00	83.00	83.00	82.82	83.15	82.91
Inference With K = 5 Shot						
3D ProtoNet	76.00	77.00	76.50	75.55	77.35	76.31
3D MatchingNet	83.00	81.00	82.00	81.90	81.75	82.19
HMCAE	86.00	88.00	87.00	88.50	87.82	86.87

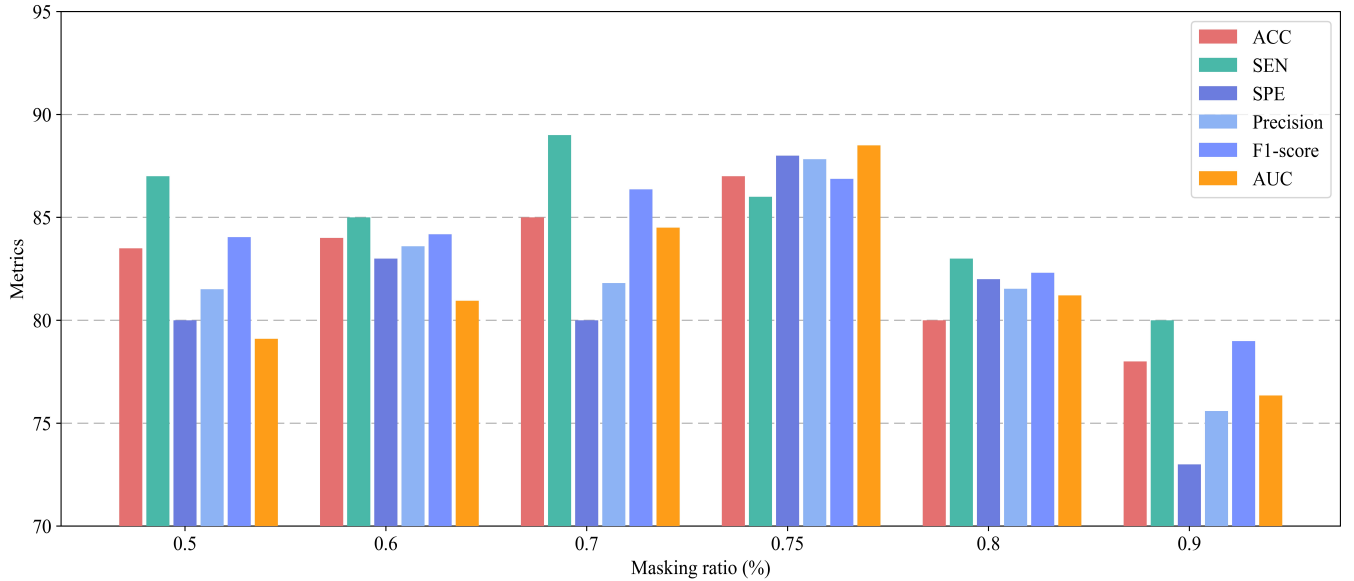


Fig. 6. The impact of different masking ratios on model performance.

and F1 score initially increased and then decreased, while specificity, precision, and AUC gradually increased. This suggests that the scale limitation of data will significantly affect the improvement of model performance in supervised learning methods. The model will learn to sacrifice sensitivity to increase specificity and improve classification accuracy. However, in medical classification tasks, model sensitivity is often the performance metric of primary interest.

Additionally, this study employed transfer learning to predict occult lymph node metastasis. Our team used a series of pretrained 3D-ResNet models from the MedicalNet project, which aggregates datasets with different modalities, target organs, and pathologies to build a relatively large dataset for extracting general medical 3D features. The results of transfer learning experiments are shown in Table I. Due to the significant differences in features between the pretraining dataset used by MedicalNet and our TJ-OLNM dataset, the performance of the pretrained models was inferior even to that of

the fully supervised learning models trained from scratch. The pretrained models had lower sensitivity and higher specificity. However, for our occult lymph node metastasis prediction task, we need higher sensitivity to reduce the false negative rate for positive patients.

In the realm of self-supervised FSL methods, our proposed HMCAE was compared with ProtoNet and MatchingNet, which are specifically designed for small sample classification tasks. To ensure fairness, all these methods utilized the self-supervised pretrained encoder of MCAE as the feature extraction framework. The experimental results demonstrate that our proposed HMCAE outperforms the baseline models in various performance indicators in both 1-shot and 5-shot scenarios. In comparison to fully supervised learning and transfer learning methods, our backbone network is simpler yet exhibits higher performance efficacy.

TABLE II

PERFORMANCE OF HMCAE MODELS WITH AND WITHOUT SELF-SUPERVISED PRETRAINING ON THE TJ-OLNM DATASET FOR OCCULT LYMPH NODE METASTASIS PREDICTION

Inference With K = 1 Shot						
Method	SEN (%)	SPE (%)	ACC (%)	AUC (%)	Precision (%)	F1 Score (%)
HMCAE w/o pretrained	80.00	70.00	75.00	77.50	72.73	76.19
HMCAE w/ pretrained	83.00	83.00	83.00	82.82	83.15	82.91
Inference With K = 5 Shot						
HMCAE w/o pretrained	80.00	75.00	77.50	81.50	76.19	78.05
HMCAE w/ pretrained	86.00	88.00	87.00	88.50	87.82	86.87

E. Ablation Study

1) *Masking Ratio*: To investigate the impact of different masking rates on model performance, we conducted self-supervised pretraining experiments with varying masking rates (50% to 90%). Fig. 6 illustrates the influence of masking rate on various model performance indicators, with the highest performance achieved at a high masking rate of 75% for most indicators (with sensitivity achieving the best results at 70%).

2) *Metric Spaces*: We also implemented our model in Euclidean space to validate the effectiveness of hyperbolic space. Table II reports the prediction results of the 3D ProtoNet in Euclidean space and the HMCAE in hyperbolic space. It shows that the HMCAE model achieves the highest scores in every evaluation metric. Specifically, the performance improvements achieved by the hyperbolic space HMCAE model compared to the Euclidean space 3D ProtoNet model are 7% in sensitivity, 10% in specificity, 7.5% in accuracy, 10.87 in AUC, 9.09% in precision, and 8.45% in F1 score for 1-shot, and 10% in sensitivity, 11% in specificity, 10.5% in accuracy, 12.95 in AUC, 10.47% in precision, and 10.56% in F1 score for 5-shot. These results vividly demonstrate that representation in the Poincaré ball has a richer embedding than that in Euclidean space and can better represent potential hierarchical clustering relationships.

3) *Self-supervised Pretraining*: To validate the advantages of the self-supervised pretraining strategy, we compared the feature extraction network with randomly initialized weights to the one that underwent self-supervised pretraining. As can be seen in Table II, the self-supervised pretraining strategy significantly improved the model's sensitivity, specificity, accuracy, AUC, precision, and F1 score. The accuracy and AUC improved by 8% and 5.32%, respectively, in 1-shot, and by 9.5% and 7%, respectively, in 5-shot. These experimental results demonstrate that self-supervised pretraining strategy helps the network to extract features closely related to occult lymph nodes.

V. CONCLUSION

In this paper, we have built the first CT image dataset for the prediction of occult lymph node metastases. Furthermore we propose a novel self-supervised hyperbolic metric few-shot learning method, OLN-Net, for metastasis prediction of occult lymph nodes in mediastinal CT scans of NSCLC patients. The OLN-Net method attempts to overcome obstacles resulting from insufficient training samples by exploiting feature information from similar CT data. OLN-Net incorporates a MCAE, a self-supervised pretraining strategy

that learns a robust image representation through mask reconstruction after eliminating background noise in CT. In the fine-tuning phase, OLN-Net uses a few-shot metric learning strategy to perform fine-tuning and reduce the disconnect between the pretrained model and real data. In addition, OLN-Net introduces hyperbolic space to represent potential hierarchical clustering relationships, which helps better learn distinguishable features for prediction. Experimental results on our collected dataset demonstrate that the proposed OLN-Net method achieves state-of-the-art results in the prediction of occult lymph node metastasis in NSCLC patients.

REFERENCES

- [1] K. C. Thandra, A. Barsouk, K. Saginala, J. S. Aluru, and A. Barsouk, "Epidemiology of lung cancer," *Contemporary Oncology*, vol. 25, pp. 45 – 52, 2021.
- [2] Y. Yun Song, "Ct radio genomics of non-small cell lung cancer using machine and deep learning," *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pp. 128–139, 2021.
- [3] X. qun He, T. Luo, X. Li, J. wen Huo, J. Gong, and Q. Li, "Clinicopathological and computed tomographic features associated with occult lymph node metastasis in patients with peripheral solid non-small cell lung cancer," *European journal of radiology*, vol. 144, p. 109981, 2021.
- [4] S. Bracci, M. Dolciemi, C. Trobiani, A. Izzo, A. Pernazza, G. d'Amati, L. Manganaro, and P. Ricci, "Quantitative ct texture analysis in predicting pd-l1 expression in locally advanced or metastatic nsccl patients," *La Radiologia Medica*, vol. 126, pp. 1425 – 1433, 2021.
- [5] W. Mu, L. Jiang, J. Zhang, Y. long Shi, J. E. Gray, I. Tunali, C. Gao, Y. Sun, J. Tian, X. Zhao, X. Sun, R. J. Gillies, and M. B. Schabath, "Non-invasive decision support for nsccl treatment using pet/ct radiomics," *Nature Communications*, vol. 11, 2020.
- [6] C. M. Walker, J. H. Chung, G. F. Abbott, B. P. Little, A. H. El-Sherief, J.-A. O. Shepard, and M. Lanuti, "Mediastinal lymph node staging: from noninvasive to surgical," *AJR. American journal of roentgenology*, vol. 199 1, pp. W54–64, 2012.
- [7] A. Szlubowski, M. Zielinski, J. Soja, A. Filarecka, S. Orzechowski, J. Pankowski, A. Obrochta, M. Jakubiak, J. Węgrzyn, and A. Ćmiel, "Accurate and safe mediastinal restaging by combined endobronchial and endoscopic ultrasound-guided needle aspiration performed by single ultrasound bronchoscope," *European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery*, vol. 46 2, pp. 262–6, 2014.
- [8] P. F. Roberts, D. M. Follette, D. von Haag, J. A. Park, P. E. Valk, T. R. Pounds, and D. M. Hopkins, "Factors associated with false-positive staging of lung cancer by positron emission tomography," *The Annals of thoracic surgery*, vol. 70 4, pp. 1154–9; discussion 1159–60, 2000.
- [9] R. Kanzaki, M. Higashiyama, A. Fujiwara, T. Tokunaga, J. Maeda, J. Okami, T. Kozuka, T. Hosoki, Y. Hasegawa, M. Takami, Y. Tomita, and K. Kodama, "Occult mediastinal lymph node metastasis in nsccl patients diagnosed as clinical n0-1 by preoperative integrated fdg-pet/ct and ct: Risk factors, pattern, and histopathological study," *Lung cancer*, vol. 71 3, pp. 333–7, 2011.
- [10] H. C. Verduzco-Aguirre, G. de Lima Lopes, and E. S.-P. de Celis, "Implementation of diagnostic resources for cancer in developing countries: a focus on pet/ct," *ecancermedicalscience*, vol. 13, 2019.

- [11] G. A. Eapen, A. M. Shah, X. Lei, C. A. Jimenez, R. C. Morice, L. Yarmus, J. J. Filner, C. Ray, G. C. Michaud, S. R. Greenhill, M. Sarkiss, R. F. Casal, D. C. Rice, and D. E. Ost, "Complications, consequences, and practice patterns of endobronchial ultrasound-guided transbronchial needle aspiration: Results of the aquire registry," *Chest*, vol. 143 4, pp. 1044–1053, 2013.
- [12] Y. Zhong, M. Yuan, T. Zhang, Y.-D. Zhang, H. Li, and T. fu Yu, "Radiomics approach to prediction of occult mediastinal lymph node metastasis of lung adenocarcinoma," *AJR. American journal of roentgenology*, vol. 211 1, pp. 109–113, 2018.
- [13] M. Cong, H. Yao, H. Liu, L. Huang, and G. Shi, "Development and evaluation of a venous computed tomography radiomics model to predict lymph node metastasis from non-small cell lung cancer," *Medicine*, vol. 99, 2020.
- [14] J. R. Ferreira, M. Koenigkam-Santos, F. E. G. Cipriano, A. T. Fabro, and P. M. A. Marques, "Radiomics-based features for pattern recognition of lung cancer histopathology and metastases," *Computer methods and programs in biomedicine*, vol. 159, pp. 23–30, 2018.
- [15] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *International Conference on Machine Learning*, 2017.
- [16] Z. Hu, Z. Gan, W. Li, J. Z. Wen, D. Zhou, and X. Wang, "Two-stage model-agnostic meta-learning with noise mechanism for one-shot imitation," *IEEE Access*, vol. 8, pp. 182 720–182 730, 2020.
- [17] J. Snell, K. Swersky, and R. S. Zemel, "Prototypical networks for few-shot learning," *ArXiv*, vol. abs/1703.05175, 2017.
- [18] O. Vinyals, C. Blundell, T. P. Lillicrap, K. Kavukcuoglu, and D. Wierstra, "Matching networks for one shot learning," in *NIPS*, 2016.
- [19] X. Liu, F. Zhou, J. Liu, and L. Jiang, "Meta-learning based prototype-relation network for few-shot classification," *Neurocomputing*, vol. 383, pp. 224–234, 2020.
- [20] A. K. Mondal, J. Dolz, and C. Desrosiers, "Few-shot 3d multi-modal medical image segmentation using generative adversarial learning," *ArXiv*, vol. abs/1810.12241, 2018.
- [21] X. Chen, J. J. Xia, D. Shen, C. Lian, L. Wang, H. H. Deng, S. Fung, D. Nie, K.-H. Thung, P.-T. Yap, and J. Gateno, "One-shot generative adversarial learning for mri segmentation of craniomaxillofacial bony structures," *IEEE Transactions on Medical Imaging*, vol. 39, pp. 787–796, 2020.
- [22] A. Paul, T. C. Shen, S. Lee, N. Balachandar, Y. Peng, Z. Lu, and R. M. Summers, "Generalized zero-shot chest x-ray diagnosis through trait-guided multi-view semantic embedding with self-training," *IEEE Transactions on Medical Imaging*, vol. 40, pp. 2642–2655, 2021.
- [23] S. Puch, I. Sánchez, and M. Rowe, "Few-shot learning with deep triplet networks for brain imaging modality recognition," in *DART/MIL3ID@MICCAI*, 2019.
- [24] A. G. Roy, S. Siddiqui, S. Pölsterl, N. Navab, and C. Wachinger, "'squeeze & excite' guided few-shot segmentation of volumetric images," *Medical image analysis*, vol. 59, p. 101587, 2019.
- [25] R. Singh, V. Bharti, V. Purohit, A. Kumar, A. K. Singh, and S. K. Singh, "Metamed: Few-shot medical image classification using gradient-based meta-learning," *Pattern Recognit.*, vol. 120, p. 108111, 2021.
- [26] D. Pathak, P. Krähenbühl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2536–2544, 2016.
- [27] C. Doersch, A. K. Gupta, and A. A. Efros, "Unsupervised visual representation learning by context prediction," *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1422–1430, 2015.
- [28] C. Doersch and A. Zisserman, "Multi-task self-supervised visual learning," *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2070–2079, 2017.
- [29] Z. Wu, Y. Xiong, S. X. Yu, and D. Lin, "Unsupervised feature learning via non-parametric instance discrimination," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3733–3742, 2018.
- [30] T. Chen, S. Kornblith, M. Norouzi, and G. E. Hinton, "A simple framework for contrastive learning of visual representations," *ArXiv*, vol. abs/2002.05709, 2020.
- [31] K. He, H. Fan, Y. Wu, S. Xie, and R. B. Girshick, "Momentum contrast for unsupervised visual representation learning," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9726–9735, 2019.
- [32] K. He, X. Chen, S. Xie, Y. Li, P. Doll'ar, and R. B. Girshick, "Masked autoencoders are scalable vision learners," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15979–15988, 2021.
- [33] L. Zhou, H. Liu, J. Bae, J. He, D. Samaras, and P. Prasanna, "Self pre-training with masked autoencoders for medical image analysis," *ArXiv*, vol. abs/2203.05573, 2022.
- [34] R. Sarkar, "Low distortion delaunay embedding of trees in hyperbolic plane," in *International Symposium Graph Drawing and Network Visualization*, 2011.
- [35] F. Sala, C. D. Sa, A. Gu, and C. Ré, "Representation tradeoffs for hyperbolic embeddings," *Proceedings of machine learning research*, vol. 80, pp. 4460–4469, 2018.
- [36] M. Nickel and D. Kiela, "Poincaré embeddings for learning hierarchical representations," in *NIPS*, 2017.
- [37] Y. Zhu, D. Zhou, J. Xiao, X. Jiang, X. Chen, and Q. Liu, "Hypertext: Endowing fasttext with hyperbolic geometry," in *Findings*, 2020.
- [38] S. Dai, Z. Gan, Y. Cheng, C. Tao, L. Carin, and J. Liu, "Apo-vae: Text generation in hyperbolic space," in *North American Chapter of the Association for Computational Linguistics*, 2020.
- [39] B. Dhingra, C. J. Shallue, M. Norouzi, A. M. Dai, and G. E. Dahl, "Embedding text in hyperbolic spaces," in *TextGraphs@NAACL-HLT*, 2018.
- [40] A. Tifrea, G. Bécigneul, and O.-E. Ganea, "Poincaré glove: Hyperbolic word embeddings," *ArXiv*, vol. abs/1810.06546, 2018.
- [41] F. López, B. Heinzerling, and M. Strube, "Fine-grained entity typing in hyperbolic space," *ArXiv*, vol. abs/1906.02505, 2019.
- [42] F. J. López and M. Strube, "A fully hyperbolic neural model for hierarchical multi-class classification," *ArXiv*, vol. abs/2010.02053, 2020.
- [43] V. Khrulkov, L. Mirvakhabova, E. Ustinova, I. Oseledets, and V. S. Lempitsky, "Hyperbolic image embeddings," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6417–6427, 2019.
- [44] D. Grattarola, L. F. Livi, and C. Alippi, "Adversarial autoencoders with constant-curvature latent manifolds," *Appl. Soft Comput.*, vol. 81, 2018.
- [45] W. Peng, J. Shi, Z. Xia, and G. Zhao, "Mix dimension in poincaré geometry for 3d skeleton-based action recognition," *Proceedings of the 28th ACM International Conference on Multimedia*, 2020.
- [46] I. Holopainen, "Riemannian geometry," *Nature*, vol. 119, pp. 117–117.
- [47] Çaglar Gülçehre, M. Denil, M. Malinowski, A. Razavi, R. Pascanu, K. M. Hermann, P. W. Battaglia, V. Bapst, D. Raposo, A. Santoro, and N. de Freitas, "Hyperbolic attention networks," *ArXiv*, vol. abs/1805.09786, 2018.
- [48] H. R. Roth, L. Lu, A. Seff, K. M. Cherry, J. Hoffman, S. Wang, J. Liu, E. B. Turkbey, and R. M. Summers, "A new 2.5d representation for lymph node detection using random sets of deep convolutional neural network observations," *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 17 Pt 1, pp. 520–7, 2014.
- [49] A. Seff, L. Lu, K. M. Cherry, H. R. Roth, J. Liu, S. Wang, J. Hoffman, E. B. Turkbey, and R. M. Summers, "2d view aggregation for lymph node detection using a shallow hierarchy of linear classifiers," *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 17 Pt 1, pp. 544–52, 2014.
- [50] A. Seff, L. Lu, A. Barbu, H. R. Roth, H.-C. Shin, and R. M. Summers, "Leveraging mid-level semantic boundary cues for automated lymph node detection," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015.