

# 基于深度语义模型的乳腺X线图像检索

邹佩, 王颖, 李洁

(西安电子科技大学电子工程学院, 西安, 710071)

**摘要:** 图像特征是基于内容的图像检索(Content-based image retrieval, CBIR)的关键, 大部分使用的手工特征难以有效地表示乳腺肿块的特征, 底层特征与高层语义之间存在语义鸿沟。为了提高CBIR的检索性能, 本文采用深度学习来提取图像的高层语义特征。由于乳腺X线图像的深度卷积特征在空间和特征维度上存在一定的冗余和噪声, 本文在词汇树和倒排文件的基础上, 对深度特征的空间和语义进行优化, 构建了两种不同的深度语义树。为了充分发挥深度卷积特征的识别能力, 根据乳腺图像深度特征的局部特性对树节点的权重进行细化, 提出了两种节点加权方法, 得到了更好的检索结果。本文从乳腺X线图像数据库(Digital database for screening mammography, DDSM)中提取了2 200个感兴趣区域(Region of interest, ROIs)作为数据集, 实验结果表明, 该方法能够有效提高感兴趣肿块区域的检索精度和分类准确率, 并且具有良好的可扩展性。

**关键词:** 乳腺肿块; 乳腺X线图像; 基于内容的图像检索; 深度学习; 深度语义树

**中图分类号:** TP391

**文献标志码:** A

## Mammogram Retrieval Based on Deep Semantic Model

ZOU Pei, WANG Ying, LI Jie

(School of Electronic Engineering, Xidian University, Xi'an, 710071, China)

**Abstract:** Image feature is the key to content-based image retrieval (CBIR). Most of the used manual features are difficult to effectively represent the features of the breast mass, and there is a semantic gap between the underlying features and the high-level semantics. In order to improve the retrieval performance of CBIR, this paper uses deep learning to extract the high-level semantic features of images. Because the deep convolution features of mammograms have some redundancies and noises in the spatial and feature dimensions, this paper optimizes the spatial and semantic features of depth features based on the vocabulary tree and inverted files, and constructs two different depth semantic trees. In order to fully exert the discriminative power of deep convolution features, the weight of tree nodes is refined according to the local characteristics of breast image depth features, and two node weighting methods are proposed to obtain better retrieval results. In this paper, 2 200 regions of interest (ROIs) are extracted from the digital database for screening mammography (DDSM) as datasets. The experimental results show that the proposed method can effectively improve the retrieval accuracy and the classification accuracy of the mass region of interest, and has good scalability.

**Key words:** breast masses; mammogram; content-based image retrieval; deep learning; depth semantic tree

## 引言

在众多威胁人类身体健康的病症中,乳腺癌是女性最常见的恶性肿瘤之一,严重威胁着女性的生命健康,且发病率呈升高的趋势,而乳腺癌的早期诊断可以大大地降低死亡率<sup>[1]</sup>。乳腺钼靶X线摄影术是早期诊断乳腺癌的首选影像学诊断工具。计算机辅助检测(Computer aided diagnosis, CAD)技术在乳腺癌的早期诊断过程中给放射科医生提供了有效的“第二意见”<sup>[2]</sup>。乳腺癌常见的症状包括肿块、微钙化点和结构畸变等类型,不同类型之间差异巨大。其中肿块作为重要的乳腺癌征象,是判定乳腺癌的重要依据。复杂的乳腺X线图像信息、肿块之间的巨大差异以及成像时的噪声干扰等情况增大了医生诊断的难度,更为乳腺肿块的检测、检索等辅助诊断任务带来了巨大的挑战。

基于内容的图像检索(Content-based image retrieval, CBIR)的CAD方法<sup>[3-7]</sup>首先促使放射科医师在查询案例中标记一个感兴趣的区域(Region of interest, ROI),将其与从先前诊断的案例中提取的数据库ROIs进行比较,并最终返回最相似的案例以及查询案例中出现肿块的可能性,帮助医生进行医学图像的诊断与治疗规划。CBIR方法的检索精度和效率在很大程度上依赖于所采用的视觉特征,而常采用的底层特征和高层语义之间存在语义鸿沟,难以有效地表示乳腺肿块的特征,从而影响检索精度。基于CBIR的CAD方法中的大多数都缺乏可伸缩性,因为它们不使用索引系统,而是将查询图像与数据集图像进行比较,这在处理时间方面是一个巨大的挑战。缺乏可伸缩性可能会阻碍这些有价值的医学图像的使用,并限制CAD诊断的准确性。为了解决这个问题,Jiang等<sup>[8]</sup>利用词汇树框架和倒排文件进行乳腺X线图像检索,词汇树的空间时间成本较低,因此具有较高的准确性和良好的可伸缩性。受Ou等<sup>[9]</sup>的启发,本文在词汇树的基础上,引入了深度学习框架,提出了一种新的深度语义树检索方法。首先,引入卷积神经网络(Convolutional neural network, CNN)模型来提取图像特征,由于乳腺X线图像的深度卷积特征在空间和特征维度上存在冗余和噪声,本文通过自编码器来改进原深度网络,对深度卷积特征进行空间和语义上的优化,获得更有效的图像表示。然后采用层次 $k$ -means聚类构建深度语义树模型,检索过程中利用深度语义树节点的特征分布,生成自适应权重来改善图像检索效果。对于基于语义优化的深度语义树,进一步提出了同心圆限制的匹配方法,对不同的图像区域分别计算相似度得分,进一步提高了图像检索精度,同时获得了较好的分类效果。对从乳腺X线图像数据库(Digital database for screening mammography, DDSM)中提取的2 200个ROIs大数据集进行的大量实验,证明了该方法的准确性和可扩展性。

## 1 相关工作

随着医学影像数量的爆发式增长,基于内容的图像检索方法在医学图像CAD系统中得到广泛应用。目前,国内外有众多学者研究了基于内容的医学图像检索(Content-based medical image retrieval, CB-MIR)在乳腺X线图像上的应用。其基本思想是提取图像的底层特征构建特征向量,以视觉相似性作为检索和匹配图像的依据。Singh等<sup>[10]</sup>利用基于共现阈值的种子区域生长算法对预处理图像进行分割,从分割区域计算出一组图像特征,包括形状特征、基于直方图的统计特征和Gabor特征等。为了选择最优特征,采用了最小冗余最大关联(minimum Redundancy maximum relevance, mRMR)特征选择方法。最后利用欧氏距离相似性测度检索相似图像。Wang等<sup>[11]</sup>提出了一种结合位置特征的相似性度量方法用于乳腺图像检索。通过提取图像的内容特征,计算内容相似度,并与位置相似度进行融合,得到图像融合的相似度。Singh等<sup>[12]</sup>使用小波基完备局部二值模式(Wavelet base complete local binary mode, W-CLBP)和 $K$ -mean聚类进行特征提取。最后利用欧氏距离相似性度量方法从距离最近的聚类中检索图像。

虽然上述方法在特定的检索过程中有一定的检索效果,但底层的视觉特征与高层语义之间存在语义鸿沟,由于乳腺肿块具有浸润性和附着力,边缘尚不清楚,图像对比度低等特点,单一的底层特征难

以表达专家对查询语义的解释。深度学习的快速发展,使它在计算机视觉领域取得了显著的成就,许多人试图利用深层网络进行图像检索。Bressan等<sup>[13]</sup>分析了深层特征在基于内容的医学图像检索中的作用,评估了几个手工特征和从深度卷积网络中提取的特征在公共乳腺图像数据集上的检索性能,证明了深度特征能在很大程度上提高相似性查询的精度。Muramatsu等<sup>[14]</sup>研究了一种使用CNN的相似性估计方案,通过CNN提取特征并与传统的3层神经网络进行相似性度量,使相似度度量与主观评价具有良好的相关性。Qayyum等<sup>[15]</sup>提出了一种基于深度卷积神经网络的医学图像分类系统,利用深度网络提取的特征和分类结果对医学图像进行检索。

基于内容的影像检索的主要方法包括基于词汇树的图像检索和基于哈希的图像检索方法。基于词汇树的图像检索方法,根据对特征的聚类情况将聚类中心视为“视觉单词”,构建树状结构来加快相似度索引的速度<sup>[8,16-21]</sup>。相比于基于穷举搜索的传统方法,基于词汇树的方法采用层次树和倒排文件,可以显著提高检索效率。而基于哈希的图像检索方法,首先根据定义的哈希函数,将高维的原始数据压缩成低维的二进制码(例如几十位到几百位二进制码),随后用汉明空间中二进制码距离相似度的计算代替高维空间中距离相似度的计算<sup>[22-24]</sup>。基于哈希的图像检索将高维特征压缩成低维度二进制码,存在多个图像对应一个二进制码的情况,其检索速度快,但获得的检索结果往往不够精确,无法返回有效的排序结果。

## 2 基于深度语义树的乳腺X线图像检索

首先介绍用于提取特征的深度网络,以及用于优化特征的自动编码器。然后在词汇树和倒排文件的基础上,构造了基于空间优化的深度语义树和基于语义优化的深度语义树,以及相应的树节点加权方法,用于计算相似度得分。

### 2.1 深度语义树

由于乳腺肿块形态复杂,边缘各异,存在病变区域和周围正常组织黏连,病灶向周围浸润,以及在成像时被其他腺体遮挡等特点,手工提取的特征不能有效地反映乳腺肿块的特征,从而影响检索的准确性。CNN模型近年来被证明能更有效地表达图像的高层语义,因此本文采用经典的AlexNet模型作为基础框架模型来获取深度卷积特征。由于普遍认为CNN模型的高层特征图更接近图像的高层语义,低层的特征图更接近图像的低级视觉表征,同时对于医学图像而言同类图像具有局部相似性,因此将网络最后1个卷积层(即第5个卷积层)特征图作为特征。

考虑到CNN卷积层特征在空间和特征维度上比较稀疏,存在信息冗余和噪声,干扰深度语义树的构建。因此,本文在深度特征提取网络的末端,加入泛化能力较强的降噪自编码器层,对特征图的空间和特征维度分别进行优化,使特征变得更加紧凑有效,获得空间或语义优化的图像表征,用于构建深度语义树,从而进一步增加深度特征的可区分性,增加检索的准确率。

获取数据库图像优化的图像特征后,为了进一步剖析深度特征之间的内在关系,将图像的深度特征量化到树结构中,对所有数据库图像的深度特征进行层次K-means聚类,构建深度语义树。深度语义树的每一个叶节点对应一个倒排文件,为图像编号的集合,记录至少有一个特征经过该节点的图像编号。而非叶节点也有对应的虚拟倒排文件,为其子孙叶节点倒排文件的串联,倒排文件的示意图如图1所示。

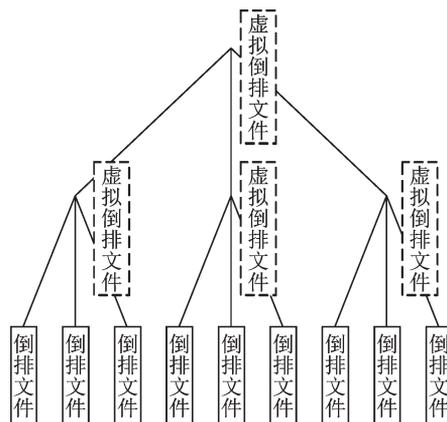


图1 倒排文件示意图

Fig.1 Reverse file schematic diagram

在深度语义树建立之后,给定一个查询图像,通过深度网络的卷积层和降噪自编码器后获得优化的深度特征。随后,将深度特征自上到下量化到深度语义树结构中。以查询图像 $q$ 的一个特征为例,迭代地选择欧式距离最近的子节点,直至达到叶节点。被选择的节点的集合被认为是特征经过这棵树的路径。查询图像 $q$ 所有特征经过的叶节点的倒排文件对应的数据库图像,即为查询图像 $q$ 的相关图像。

### 2.2 基于空间优化的深度语义树检索方法

深度特征在空间和语义特征维度上存在冗余和噪声,使用降噪自编码器对特征分别进行空间和语义的优化,得到两种不同的深度特征,构建深度语义树用于图像检索。基于空间优化的深度语义树检索方法的框架如图2所示。

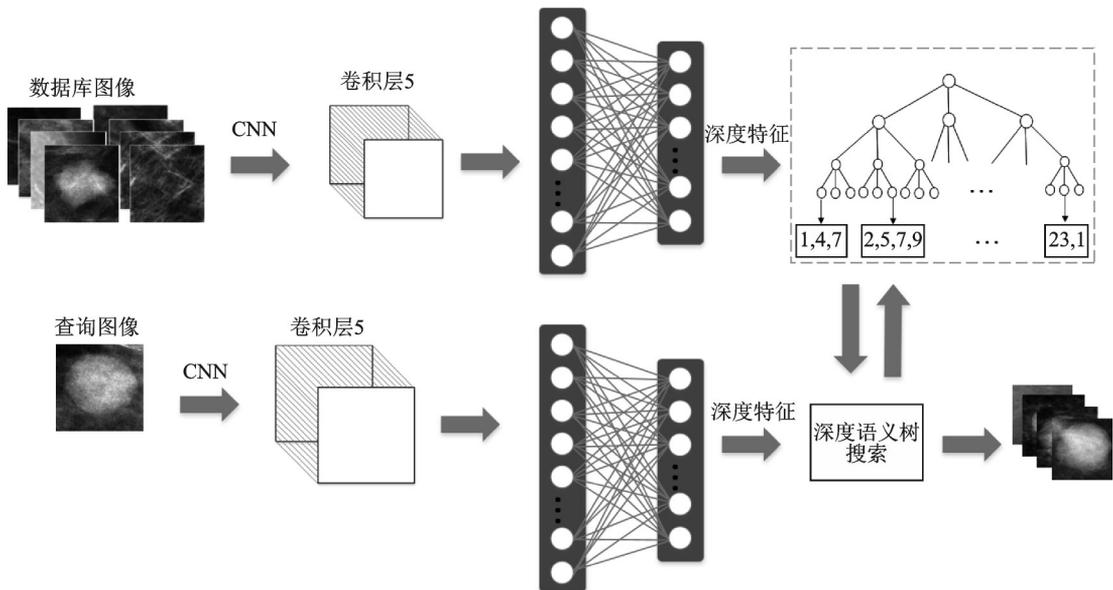


图2 基于空间优化的深度语义树检索总体框架

Fig.2 Overall block diagram of deep semantic tree retrieval based on spatial optimization

使用 AlexNet 网络提取 256 个大小为  $13 \times 13$  的特征图,每张图像对应的特征图可以看作 169 个 256 维的稠密特征。经过 ReLU 激活后的特征图存在大量值为 0 的点,在空间上存在大量冗余,本文采用的乳腺 X 线图像存在噪声、血管等干扰因素,并且对肿块图像,存在肿块背景的信息冗余,对于正常组织图像,存在纹理大量重复的信息冗余,因此对特征图空间上的优化变得十分必要。以  $F$  个  $D$  维的稠密特征作为降噪自编码器的输入,隐层输出为  $F'$  个  $D$  维特征,其中,  $F' < F$ ,从而达到特征图空间优化的目的。用降噪自编码器进行特征图空间优化的示意图如图 3 所示。

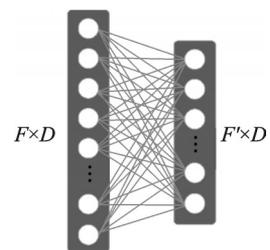


图3 用自编码器进行特征图空间优化的示意图

Fig.3 Schematic diagram of feature space optimization with a self-encoder

在得到数据库图像空间优化的图像特征后,按照前一节所述方法构建语义树得到查询图像的相关图像,并根据相似度评分测量查询图像与其相关图像的相似度。首先需要计算查询图像 $q$ 和其相关图像 $d$ 的每一对路径之间的相似度,查询图像 $q$ 的第 $i$ 条路径和其相关图像 $d$ 的第 $j$ 条路径之间的相似度得

分计算方式为

$$\text{Sim}_{p_0}(P_{q,i}, P_{d,j}) = \sum_{l=1}^L \omega_0(v_{q,i}^l) \times \delta(v_{q,i}^l, v_{d,j}^l) \quad (1)$$

式中: $\omega$ 为加权函数, $\delta$ 为克罗内克函数, $\delta(a, b) = 1$  if  $a = b$  and  $\delta(a, b) = 0$ ,节点初始权重采用逆向文件频率(Inverse document frequency, IDF)方法加权,得到

$$\omega_0(v) = \text{IDF}(v) = \log \frac{N}{N_v} \quad (2)$$

式中: $N$ 为数据库图像个数, $N_v$ 为至少有一个特征经过节点 $v$ 的数据库图像个数。将查询图像 $q$ 和相关图像 $d$ 之间所有路径的相似度相加,则可以得到查询图像 $q$ 和相关图像 $d$ 的相似度得分

$$\text{Sim}_0(q, d) = \sum_{i=1}^m \sum_{j=1}^m \text{Sim}_{p_0}(P_{q,i}, P_{d,j}) \quad (3)$$

用以上公式计算查询图像 $q$ 和其所有相关图像的相似度得分,选择相似度最高的部分相关图像作为检索结果。在检索阶段结束后,查询图像可以根据得到的相关图像进行分类。取相似度得分排名前 $t$ 的相关图像,将相关图像的相似度得分按照类别分别相加,将相似度得分累加值最高的那一类作为查询图像 $q$ 的分类结果,即有

$$CL(q) = \arg \max_{CL} \sum_{d_i \in CL} \text{Sim}(q, d_i) \quad (4)$$

式中 $d_i$ 为查询图像 $q$ 的第 $i$ 个检索结果。

每个节点对于整个语义树的重要程度不同,因此需要对每个节点赋予不同的权重。由于每个深度特征为图像对应区域的一种特征表示,对于乳腺图像,肿块图像通常包含了部分背景(即正常组织)及其他和正常组织具有相似特点的区域,这部分区域提取的特征极有可能和正常组织图像的特征更相似。因此在构建深度语义树的过程中,容易被分类到正常组织图像较多的节点,这部分特征对于检索不具备有效性,是相似度得分计算过程中的干扰项。因此,应在计算路径相似度时加入自适应权重,相关图像所属图像种类在该节点占的比重越大,则认为该节点对于该类图像越可靠,其权重越大。自适应权重的计算公式为

$$\omega(v) = \begin{cases} \frac{N_{mv}}{N_v} \times \omega_0(v) & d \in \text{mass} \\ \frac{N_{nv}}{N_v} \times \omega_0(v) & d \in \text{normal} \end{cases} = \begin{cases} \frac{N_{mv}}{N_v} \times \log \frac{N}{N_v} & d \in \text{mass} \\ \frac{N_{nv}}{N_v} \times \log \frac{N}{N_v} & d \in \text{normal} \end{cases} \quad (5)$$

式中: $N_{mv}$ 和 $N_{nv}$ 分别表示经过节点 $v$ 的肿块图像个数和正常组织图像个数。 $N_v$ 表示该节点的图像个数。因此,查询图像 $q$ 的第 $i$ 条路径和其相关图像 $d$ 的第 $j$ 条路径之间的相似度得分计算时采用新的权重,相似度计算方式为

$$\text{Sim}_p(P_{q,i}, P_{d,j}) = \sum_{l=1}^L \omega(v_{q,i}^l) \times \delta(v_{q,i}^l, v_{d,j}^l) \quad (6)$$

### 2.3 基于语义优化的深度语义树检索方法

经过不同的卷积模板,卷积层的每个特征图是对原始图像不同特性的反映。在卷积层各个维度的特征图中,存在很多零矩阵和非零元素极少的矩阵,因此有些特征图对于目标表示而言效果较弱。在特征维度上存在一定的信息冗余,从而干扰深度语义树的构建,因此对特征维度的优化十分必要。以 $F$ 个 $D$ 维的稠密特征作为降噪自编码器的输入,隐层输出为 $F$ 个 $D'$ 维特征,其中, $D' < D$ ,从而达到特征图语义优化的目的。

用降噪编码器进行特征图语义优化的示意图如图4所示。在得到数据库图像语义优化的图像特征后,按照前一节所述方法构建语义树,得到查询图像的相关图像之后,在沿用自适应权重方法的同时,考虑肿块通常在图像的中心位置,图像四周位置为肿块的背景区域,因此不同部位的特征重要程度不同。在深度语义树的检索过程中,容易造成肿块图像背景和正常组织图像中心位置之间的无效比对。为了使特征路径的比对更有意义,本文提出基于同心圆限制的匹配方法,按照图5所示划分同心圆,在比较两个图像相似度时,改变原本一幅图像每条路径和另一幅图像每条路径进行对比的方式,只比较这两幅图像对应区域的所有路径,将每个区域的得分加权相加得到两个图像之间的相似度。因此查询图像 $q$ 和其相关图像 $d$ 的相似度得分计算方式为

$$\text{Sim}(q, d) = \sum_{b=1}^B \lambda_b \sum_{i \in B, j \in B} \text{Sim}_p(P_{q,i}, P_{d,j}) = \sum_{b=1}^B \lambda_b \sum_{i \in c, j \in c, l=1}^L \omega(v_{q,i}^l) \times \delta(v_{q,i}^l, v_{d,j}^l) \quad (7)$$

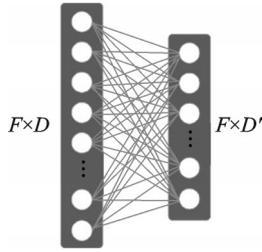


图4 用自编码器进行语义优化的示意图

Fig.4 Schematic diagram of semantic optimization with a self-encoder

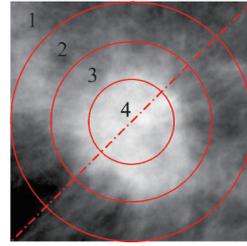


图5 同心圆划分示意图

Fig.5 Diagram of concentric circle division

### 3 实验结果与分析

本文将从检索性能和分类性能两方面对提出方法的性能进行评估。首先描述了实验设置,包括数据集、评价方案和比较方法,之后给出了实验结果并进行了分析。

#### 3.1 实验设置

本文的实验数据集基于乳腺X线图像数据库构建。提取了由经验丰富的放射科医生注释的以肿块(1 100)为中心的乳腺X线ROI,同时提取1 100个正常区域。最后在上述ROIs中,随机选取1 000个肿块ROIs和1 000个正常组织ROIs(2 000个ROIs)作为数据库,构建深度语义树。随机选择200个ROIs作为查询图像,其中一半为肿块区域另一半为正常区域,每次返回5个图像作为检索结果。为保证实验结果的可靠性,每组对比实验重复100次。在评价检索性能时,本文采用检索精度(Precision)作为算法检索性能的评价指标,其计算公式为

$$\text{Precision} = \frac{TP}{TP + FP} \quad (8)$$

式中: $TP$ 为返回的检索结果中真阳的个数, $FP$ 为返回的检索结果中假阳的个数,检索精度也可以解释为检索结果中符合要求的图像数和检索返回的图像总数之比。

在评价分类性能时采用分类准确率(Accuracy)以及受试者工作特征(Receiver operating characteristic, ROC)曲线进行算法性能评估。分类准确率反映了分类正确的结果占所有分类结果的比例,计算公式为

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (9)$$

式中:  $TP$  为分类结果中真阳结果的个数,  $FP$  表示分类结果中假阳结果的个数,  $TN$  表示分类结果中真阴结果的个数,  $FN$  表示分类结果中假阴结果的个数。

对于检索性能, 将本文的方法与其他现有的检索方法进行比较, “BoW” 表示提取图像的尺度不变特征变换 (Scale invariant feature transform, SIFT) 特征后用 BoW 构建词典, 并将特征量化到词典中进行检索的方法<sup>[25]</sup>, 是比较经典的图像检索方法, 在实验过程中选取聚类中心数为 1 000; “VOC-TREE” 表示 Jiang 等<sup>[8]</sup>提出的词汇树进行乳腺 X 线图像检索的方法; “C-AGH-ITQ” 表示 Liu 等<sup>[22]</sup>提出的用于乳腺 X 线图像检索的融合 SIFT 特征和 GIST 特征的无监督哈希方法; “CNN-KSH” 表示提取 AlexNet 模型卷积层特征, 进行核哈希的检索方法<sup>[21]</sup>; “SP-OP-TREE” 表示本文提出的基于空间优化的深度语义树检索方法, “SE-OP-TREE” 表示本文提出的基于语义优化的深度语义树检索方法<sup>[24]</sup>。在分类性能方面, 也将本文的方法与现有分类方法进行了比较, “CNN-SVM” 表示将 AlexNet 第 5 个卷积层接 SVM 的分类方法, “CNN-Forest” 表示将 AlexNet 第 5 个卷积层接随机森林的方法。所有实验都在 MATLAB 2017a 环境下进行, 运行环境为 Intel Core i7 4770K 处理器, 16 GB 内存, Windows 10 操作系统。

### 3.2 实验结果与分析

为了定性证明用降噪自编码器进行进一步处理之后的特征比卷积层特征具有更好的可区分性, 采用  $t$ -SNE ( $t$ -distributed stochastic neighbor embedding,  $t$ -SNE) 算法<sup>[26]</sup>对特征进行可视化。分别将最后一个卷积层的输出和降噪自编码器的输出作为特征进行比较, 由于获得的特征均为图像的局部表征, 为便于可视化显示, 将每张图像的特征分别串联成一个向量, 随后分别将两种特征用  $t$ -SNE 非线性降维至二维平面, 使每个数据对应二维的一个位置, 从而实现可视化。空间优化的特征可视化图如图 6 所示, 语义优化的特征可视化图如图 7 所示。两图中标记为 1 的红色圆点为肿块图像的特征, 标记为 0 的蓝色圆点为正常组织图像的特征。由图中数据分布可知, 用降噪自编码器进行空间优化和语义优化之后的深度特征可区分性更强。

为了验证本文提出的权重方法的有效性, 本文分别比较了基于空间优化的深度语义树和基于语义优化的深度语义树各部分对于检索结果的影响, 比较了不同权重时的检索精度。基于空间优化的深度语义树算法各部分对实验结果的影响如表 1 所示。“CNN” 表示 CNN 卷积层直接接入树结构, 相似度计算采用初始权重; “CNN-SPAUTO” 表示 CNN 卷积层后连接自编码器进行空间优化之后接入树结构, 相似度计算采用初始权重; “SP-OP-TREE” 表示基于空间优化的深度语义树算法。由“CNN” 和“CNN-SPAUTO” 的实验结果可知, 用降噪自编码器进行空间优化能够有效地提高图像的检索精度; 由“SP-OP-TREE” 和“CNN-SPAUTO” 的实验结果可知, 用自适应权重方法动态调整树节点的权重, 能够

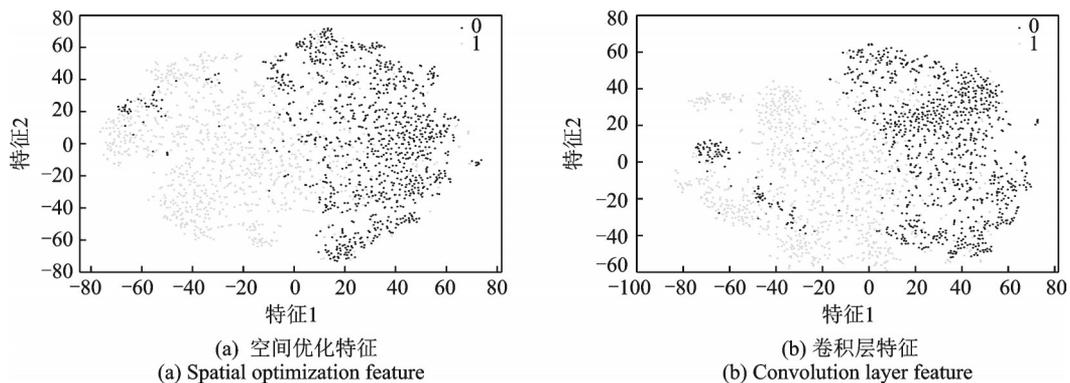


图6 空间优化的特征可视化结果

Fig.6 Feature visualization of spatial optimization

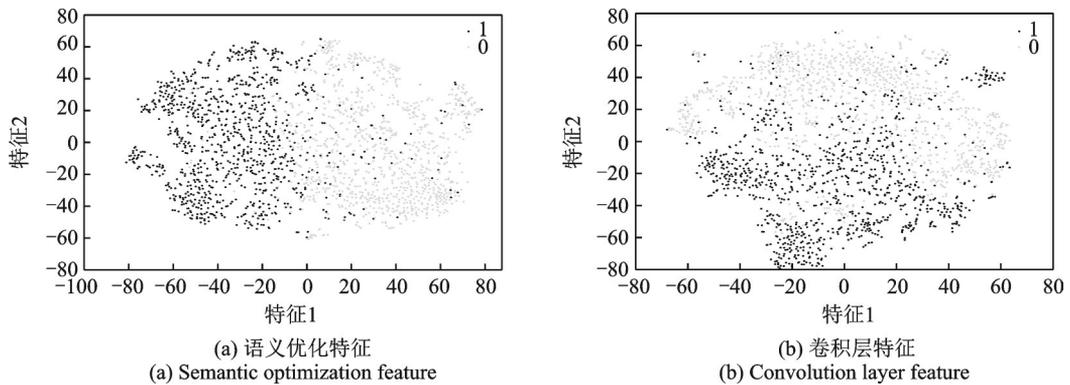


图7 语义优化的特征可视化结果

Fig.7 Feature visualization of semantic optimization

表1 基于空间优化的深度语义树算法各部分对实验结果的影响

Table 1 Influence of each part of the deep semantic tree algorithm based on spatial optimization on the experimental results %

方法	肿块检索精度	正常组织检索精度	平均精度
CNN	83.9	88.3	86.1
CNN-SPAUTO	86.0	92.1	89.1
SP-OP-TREE	87.3	93.6	90.5

进一步提高检索效果。

基于语义优化的深度语义树算法各部分对实验结果的影响如表2所示。“CNN-SEAUTO”表示CNN卷积层接自编码器进行语义优化后接入树结构,相似度计算采用初始权重;“CNN-WE”表示CNN卷积层直接接入树结构,采用自适应权重方法进行相似度计算;“CNN-CC”表示CNN卷积层直接接入树结构,采用基于同心圆限制的匹配方法进行相似度计算,由表中实验结果可知,各项改进方法对实验效果都有一定的提升作用。其中,自编码器语义优化和自适应权重分别将肿块检索精度提高了2.3%和2.9%,较大幅度地提高了肿块检索精度,基于同心圆的限制匹配方法在肿块检索精度和正常组织检索精度上均提升了1.9%。采用了多项改进方法的基于深度语义优化的深度语义树检索方法,在检索精度上获得了极大的提升,其中,肿块检索精度提升了9.7%,正常组织检索精度提高了3.3%,平均精度提高了6.5%,对于乳腺X线图像而言,肿块检索精度的提高能减少医生的误检率。

按照3.1节的检索精度公式可计算不同检索方法的检索精度,如表3所示。由实验结果可知,相对于其他无监督检索方法(即表格前3种方法),本文提出的方法不管是在肿块检索精度还是在正常

表2 基于语义优化的深度语义树算法各部分对实验结果的影响

Table 2 Influence of each part of the deep semantic tree algorithm based on semantic optimization on the experimental results %

方法	肿块检索精度	正常组织检索精度	平均精度
CNN	83.9	88.3	86.1
CNN-SEAUTO	86.2	89.9	88.1
CNN-WE	86.8	88.6	87.7
CNN-CC	85.8	90.2	88.0
SE-OP-TREE	93.6	91.6	92.6

表3 不同检索方法的检索精度对比

Table 3 The retrieval precision of different methods

%

方法	肿块检索精度	正常组织检索精度	平均精度
BoW	76.0	71.9	74.0
VOC-TREE	80.6	74.4	77.5
C-AGH-ITQ	82.9	89.2	86.1
CNN-KSH	90.6	89.2	89.9
SP-OP-TREE	87.3	93.6	90.5
SE-OP-TREE	93.6	91.6	92.6

组织检索精度上都明显优于其他方法。其中,SP-OP-TREE的肿块检索精度比BoW高11.3%,比词汇树高6.7%,比“C-AGH-ITQ”高4.4%,正常组织检索精度甚至比BoW高了21.7%,平均精度也远超其他无监督算法,但是对于采用CNN特征的有监督算法“CNN-KSH”,SP-OP-TREE的肿块检索精度略低,而正常组织检索精度较高,平均精度仍略高于“CNN-KSH”。但由于KSH算法的编码长度有限,其返回的图像序列大致相同,对放射科医生而言参考价值不大,而SP-OP-TREE方法返回的序列各异,是根据肿块和正常组织本身特点找到的最相近的图像,并且总体精度更高,因此更具有参考意义。SE-OP-TREE的肿块检索精度同样比现有检索方法具有更好的性能,平均精度也远超其他无监督算法,对于有监督算法“CNN-KSH”,SE-OP-TREE在肿块和正常组织精度上都高于“CNN-KSH”。由于在相似度计算过程中加入了同心圆限制,SE-OP-TREE的检索精度总体优于“SP-OP-TREE”,SE-OP-TREE的肿块检索精度显著提升了6.3%,正常组织检索精度略低于“SP-OP-TREE”,平均准确率比“SP-OP-TREE”高2.1%。综上分析,SE-OP-TREE算法能够更好地挖掘图像语义特征,得到更好的检索效果。

为了进一步验证本文所提出方法的有效性,按照3.1节的分类准确率公式,计算了不同分类方法的分类准确率,如表4所示。由结果可知,本文提出的方法整体准确率优于其他方法,并且肿块分类准确率和正常组织分类准确率都高于其他方法。对于肿块分类准确率,SP-OP-TREE明显优于“CNN-FOREST”,略优于“CNN-SVM”,SE-OP-TREE比表中其他方法准确率高6.4%~8.1%;对于正常组织,SP-OP-TREE和SE-OP-TREE都显著地优于“VOC-TREE”和“CNN-SVM”(比两者高10%有余),SP-OP-TREE比“CNN-FOREST”提高了2.7%。与SP-OP-TREE方法相比,SE-OP-TREE方法在正常组织准确率上略低,但平均精度更高。各种方法的ROC曲线如图8所示。

表4 不同方法的分类准确率对比

Table 4 Classification accuracy of different methods

%

方法	肿块 准确率	正常组织 准确率	平均 准确率
VOC-TREE	87.1	78.4	83.8
CNN-SVM	87.9	82.9	85.4
CNN-FOREST	86.5	90.2	88.4
SP-OP-TREE	88.2	92.9	90.6
SE-OP-TREE	94.6	91.7	93.2

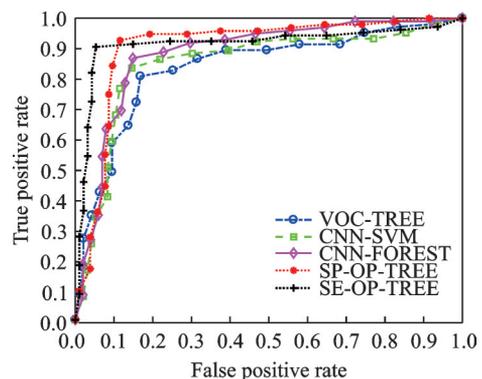


图8 不同方法的ROC曲线对比图

Fig.8 Comparison of ROC curves of different methods

## 4 结束语

本文对基于内容的医学影像检索进行了深入研究,在词汇树和反向文件的基础上,引入深度卷积特征来有效表示乳腺X线图像,并引入降噪自编码器来改进深度网络特征的获取和优化过程,对特征的空间和语义分别进行优化,得到基于空间优化的深度语义树和基于语义优化的深度语义树检索方法,为了充分发挥深度卷积特征的识别能力,根据乳腺图像深度特征的局部特性,分别对两棵树的树节点权重进行细化,得到了更好的检索结果。从乳腺X线图像数据库 DDSM 中提取了 2 200 个 ROIs 作为数据集,实验结果表明,该方法能够有效提高感兴趣肿块区域的检索精度和分类准确率,并且具有良好的可扩展性。深度语义树能够有效挖掘特征之间的内在关系,但在很大程度上仍然属于无监督模型,如何有效利用类标等监督信息对树结构进行进一步的调整,从而达到更好的检索效果也是今后的一个研究方向。

### 参考文献:

- [1] 胥林花,华琳月. 乳腺癌早期治愈率可达 95%[N]. 南京晨报, 2015-7-21(A31).
- [2] LINDA J, BURHEME W, SUSAN A, et al. Potential contribution of computer-aided detection to the sensitivity of screening mammography[J]. *Radiology*, 2000, 215(2): 554-562.
- [3] RUI Y, HUANG T S, CHANG S F. Image retrieval: Current techniques, promising directions, and open issues[J]. *Journal of Visual Communication and Image Representation*, 1999, 10(1): 39-62.
- [4] SMEULDERS A W M, WORRING M, SANTINI S, et al. Content-based image retrieval at the end of the early years[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(12): 1380.
- [5] LONG F, ZHANG H, FENG D D. Fundamentals of content-based image retrieval[M]. *Multimedia Information Retrieval and Management*. Berlin, Heidelberg: Springer, 2003: 1-26.
- [6] LIU Y, ZHANG D, LU G, et al. A survey of content-based image retrieval with high-level semantics[J]. *Pattern Recognition*, 2007, 40(1): 262-282.
- [7] DATTA R, JOSHI D, LI J, et al. Image retrieval: Ideas, influences, and trends of the new age[J]. *ACM Computing Surveys (Csur)*, 2008, 40(2): 1-60.
- [8] JIANG M, ZHANG S, LI H, et al. Computer-aided diagnosis of mammographic masses using scalable image retrieval[J]. *IEEE Transactions on Biomedical Engineering*, 2014, 62(2): 783-792.
- [9] OU X Y, LING H F, YAN L Y, et al. Convolutional neural codes for image retrieval[C]//*Proceedings of Signal and Information Processing Association Summit and Conference*. [S.l.]: IEEE, 2015: 1-10.
- [10] SINGH V P, SRIVASTAVA S, SRIVASTAVA R. Automated and effective content-based image retrieval for digital mammography[J]. *Journal of X-ray science and technology*, 2018, 26(1): 29-49.
- [11] WANG Z, XIN J, HUANG Y, et al. A similarity measure method combining location feature for mammogram retrieval[J]. *Journal of X-ray science and technology*, 2018, 26(4): 553-571.
- [12] SINGH V P, SRIVASTAVA R. Content-based mammogram retrieval using wavelet based complete-LBP and K-means clustering for the diagnosis of breast cancer[J]. *International Journal of Hybrid Intelligent Systems*, 2017, 14(1/2): 31-39.
- [13] BRESSAN R S, ALVES D H A, VALERIO L M, et al. DOCToR: The role of deep features in content-based mammographic image retrieval[C]//*Proceedings of 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*. [S.l.]: IEEE, 2018: 158-163.
- [14] MURAMATSU C, HIGUCHI S, MORITA T, et al. Retrieval of reference images of breast masses on mammograms by similarity space modeling[J]. *14th International Workshop on Breast Imaging (IWBI 2018)*, International Society for Optics and Photonics, 2018, 10718: 1071809.
- [15] QAYYUM A, ANWAR S M, AWAIS M, et al. Medical image retrieval using deep convolutional neural network[J]. *Neurocomputing*, 2017, 266: 8-20.
- [16] WANG X, YANG M, COUR T, et al. Contextual weighting for vocabulary tree based image retrieval[C]//*Proceedings of*

- 2011 International Conference on Computer Vision.[S.l.]: IEEE, 2011: 209-216.
- [17] LI Q, XU R, ZHAO H, et al. Computer-aided diagnosis of mammographic masses using local geometric constraint image retrieval[J]. *Optik*, 2018, 171: 754-767.
- [18] WANG J, XIAO J, LIN W, et al. Discriminative and generative vocabulary tree: with application to vein image authentication and recognition[J]. *Image and Vision Computing*, 2015, 34: 51-62.
- [19] WANG X, YANG M, COUR T, et al. Contextual weighting for vocabulary tree based image retrieval[C]// Proc IEEE Int Conf Comput Vis. [S.l.]:[s.n.], 2011: 209-216.
- [20] NISTER D, STEWENIUS H. Scalable recognition with a vocabulary tree[C]//Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06).[S.l.]: IEEE, 2006, 2: 2161-2168.
- [21] ZHANG S, YANG M, COUR T, et al. Query specific fusion for image retrieval[C]//Proceedings of European Conference on Computer Vision. Heidelberg, Berlin: Springer, 2012: 660-673.
- [22] LIU J, ZHANG S, LIU W, et al. Scalable mammogram retrieval using composite anchor graph hashing with iterative quantization[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016, 27(11): 2450-2460.
- [23] LIU J, ZHANG S, LIU W, et al. Scalable mammogram retrieval using anchor graph hashing[C]//Proceedings of 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI).[S.l.]: IEEE, 2014: 898-901.
- [24] ZHANG X, LIU W, DUNDAR M, et al. Towards large-scale histopathological image analysis: Hashing-based image retrieval[J]. *IEEE Transactions on Medical Imaging*, 2014, 34(2): 496-506.
- [25] WANG J, LI Y, ZHANG Y, et al. Bag-of-features based medical image retrieval via multiple assignment and visual words weighting[J]. *IEEE Transactions on Medical Imaging*, 2011, 30 (11): 1996-2011.
- [26] MAATEN L, HINTON G. Visualizing data using t-SNE[J]. *Journal of Machine Learning Research*, 2008, 9: 2579-2605.

#### 作者简介:



邹佩(1993-),女,硕士研究生,研究方向:智能信息处理、影像学处理与分析,E-mail: zoupei\_93@yeah.net。



王颖(1981-),女,博士,副教授,研究方向:影像处理与分析、模式识别。



李洁(1972-),女,博士,教授,研究方向:图像处理、智慧交通。

(编辑:刘彦东)