

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2023)02-0601-12

论文引用格式: Guan A N, Liu L, Fu X D, Liu L J and Huang Q S. 2023. A deep hash retrieval for large-scale chest radiography images. Journal of Image and Graphics, 28(02): 0601-0612 (管安娜, 刘骊, 付晓东, 刘利军, 黄青松. 2023. 面向大规模胸片图像的深度哈希检索. 中国图象图形学报, 28(02): 0601-0612) [DOI: 10.11834/jig.211012]

面向大规模胸片图像的深度哈希检索

管安娜¹, 刘骊^{1,2*}, 付晓东^{1,2}, 刘利军^{1,2}, 黄青松^{1,2}

1. 昆明理工大学信息工程与自动化学院, 昆明 650500; 2. 云南省计算机技术应用重点实验室, 昆明 650500

摘要: **目的** 医学图像检索在疾病诊断、医疗教学和辅助症状参考中发挥了重要作用, 但由于医学图像类间相似度高、病灶易遗漏以及数据量较大等问题, 使得现有哈希方法对病灶区域特征的关注较少, 图像检索准确率较低。对此, 本文以胸部 X-ray 图像为例, 提出一种面向大规模胸片图像的深度哈希检索网络。**方法** 在特征学习部分, 首先采用 ResNet-50 作为主干网络对输入图像进行特征提取得到初步特征, 将该特征进行细化后获得全局特征; 同时将初步特征输入构建的空间注意模块, 该注意模块结合了 3 个描述符用于聚焦胸片图像中的显著区域, 将该模块的输出进行细化得到局部特征; 最后融合全局特征与局部特征用于后续哈希码优化。在哈希码优化部分, 使用定义的二值交叉熵损失、对比损失和正则化损失的联合函数进行优化学习, 生成高质量的哈希码用于图像检索。**结果** 为了验证方法的有效性, 在公开的 ChestX-ray8 和 CheXpert 数据集上进行对比实验。结果显示, 构建空间注意模块有助于关注病灶区域, 定义特征融合模块有效避免了信息的遗漏, 联合 3 个损失函数进行优化可以获得高质量哈希码。与当前先进的医学图像检索方法比较, 本文方法能够有效提高医学图像检索的准确率, 在两个数据集上的检索平均精度分别提高了约 6% 和 5%。**结论** 在大规模胸片图像检索中, 本文提出的深度哈希检索方法能够有效关注病灶区域, 提高胸片图像检索的准确率。

关键词: 医学图像检索; 注意力机制; 特征融合; 深度哈希 (DH)

A deep hash retrieval for large-scale chest radiography images

Guan Anna¹, Liu Li^{1,2*}, Fu Xiaodong^{1,2}, Liu Lijun^{1,2}, Huang Qingsong^{1,2}

1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China;

2. Computer Technology Application Key Laboratory of Yunnan Province, Kunming 650500, China

Abstract: **Objective** Big medical data is mainly concerned of data storage-related like electronic healthy profiles, medical image and genetic information. It is essential to process large-scale medical image data efficiently. For large-scale retrieval tasks, deep hashing methods can be used to optimize traditional retrieval methods. To improve its retrieval efficiency, the potential ability is developed to map the high-dimensional features of an image into the binary space, generate low-dimensional binary encoded features, and avoid dimensional catastrophe problem. Hash-depth methods are divided into two data categories of those are independent data and dependent data. Although the deep hashing method has great advantages in large-scale image retrieval, the challenges are still to be resolved for the features loss issues of key areas of the lesions like

收稿日期: 2021-10-28; 修回日期: 2022-01-19; 预印本日期: 2022-01-26

* 通信作者: 刘骊 kmust_mary@163.com

基金项目: 国家自然科学基金项目 (62262036, 61862036, 61962030); 云南省中青年学术和技术带头人后备人才培养计划项目 (202005AC160036)

Supported by: National Natural Science Foundation of China (62262036, 61862036, 61962030); Yunnan Provincial Foundation for Leaders of Disciplines in Science and Technology (202005AC160036)

redundant lesions, high noise, and small targets. So, we develop a deep hash retrieval network for large-scale human chest-related X-ray images. **Method** For the feature learning part: to obtain their initial features, the ResNet-50 is used as the backbone network and the input image is subjected to feature extraction. To obtain global features, a feature-refined block is followed. Here, the feature-refined block is structured via the residual block and the average pooling layer. To obtain the detailed focal regions, we design a spatial attention module in related to three descriptors: 1) maximum element along the channel axis, 2) average element, and 3) maximum pooling. In addition, to obtain a feature focusing on the prominent region, the key features are input into the spatial attention module, and then local features are obtained in terms of feature-refined block. First, the resulting global and local features are integrated seamlessly by dimension. Next, to optimize hash codes, the cascade layer is connected to the fully-connected layer. For the part of hash code optimization: in order to obtain high quality hash codes and improve the quality of sorting results, a joint loss function is used to define the target error. To generate a more discriminative hash code, we leverage the label information and semantic features of the image in related to the losses of contrast, regularization and cross entropy. Finally, the searching results are calculated in terms of the similarity metric. **Result** The comparative experiments are carried out on two different datasets of those are ChestX-ray8 and CheXpert. Our analysis is compared to other five classical generic hashing methods for the same task, including deep hashing methods and shallow hashing methods. Among them, the deep hashing methods are based on deep hashing (DH), deep supervised hashing (DSH), and attention-based triplet hashing (ATH), and the shallow hashing methods are based on semi-supervised hashing (SSH) and iterative quantization (ITQ). The normalized discounted cumulative gain (nDCG@100) and mean average precision (mAP) are used as evaluation metrics. The experimental results show that the retrieval performance of our method has some optimal value in comparison with deep learning-relevant methods. For the ChestX-ray8 dataset, the mAP is increased by about 6% and the nDCG@100 is improved by 4%. For the CheXpert database, the mAP is higher by about 5% and the nDCG@100 is improved by 3%. **Conclusion** To deal with the problem that the existing hashing methods pay less attention to salient region features, we demonstrate a deep hash retrieval network for large-scale chest X-ray images for large-scale human chest-relevant radiographic image retrieval. To improve the accuracy and ranking quality of image retrieval, this deep hash retrieval method is proposed and be focused on the lesion region effectively. It is beneficial to clarify focal area information and reveal the attention-less problem to salient areas in terms of a spatial attention module-constructed. The feature fusion-defined module can be used resolve the problem of information loss effectively. We use three loss functions to make the real value output more similar to the binary hash code, which can optimize the sorting quality problem of the retrieval results. It is possible to adjust the order of network composition for the regions of interest (RoI)-concerned. The loss function can be optimized the existing hashing method. It is potential to distinguish small sample images further.

Key words: medical image retrieval; attention mechanism; feature fusion; deep hashing (DH)

0 引言

随着医院 PACS(the picture archival and communication system)的广泛应用,医学图像的数据量迅速增加,高效处理大规模的医学图像数据成为辅助医疗诊断的有效途径之一。通过有效的医学图像检索不仅能够帮助临床专家对疾病的预测和诊断进行决策,而且可以利用检索得到以往类似图像和病例信息,以实现医疗教学;另外,通过精准的医疗图像检索可提供与患者相似的病例图像,作为辅助病症参考。基于文本和基于内容的图像检索方法在医学

图像检索中有许多应用(Hwang 等,2012)。基于文本的图像检索是通过人工注释的文本描述来检索图像,既耗时又费力,而且用有限的文字难以描述医学图像的内容。相比之下,基于内容的图像检索使用源自图像的特征,例如纹理、颜色和形状,不需要费时和主观的人工标记。例如,吴梦麟等人(2013)针对医学图像底层特征难以描述语义信息的问题,采用图半监督学习提取语义相似度,并结合底层特征和语义特征进行相似性度量。

区别于通用图像,医学图像侧重于提取病灶区域信息来检索具有相似病例的图像(Li 等,2018),其特殊性在于:1)大多数是灰度图,噪声较大,某些

传统特征并不适用医学图像检索;2)由于成像方式原因,不具备病理意义的区域可能会模糊关键区域之间的差距;3)临床上有用的信息大多高度局限于图像的小区域,存在极大的多样性和复杂性,单独提取它们的全局特征通常不能很好地表示医学图像。本文以胸片图像为例,胸片图像就是胸部X光片,为常规体检项目之一。其不仅具有医学图像的特殊性,也有其独特的地方。速度和经济的优势使胸部X光片成为胸部检查的首选,因此具有较大的图像数据库。此外,胸部X光可以清楚地发现肺部的大体病变,例如肺炎和肿块,所以在临床护理和流行病学研究中发挥着重要作用。

Silva 等人(2020)提出一种基于可解释性显著图的方法来检索胸片图像,学习显著图得到深层特征表示,通过欧几里得排序进行检索。Kashif 等人(2020)从图像中提取特征并计算语义和视觉相似度,使用无向图进行最短路径算法来检索相似图像。然而,以上方法的检索性能在大数据库上通常受到限制。传统的图像检索方法已经难以高效处理海量数据,将特征哈希与深度学习相结合的深度哈希技术已成为图像检索的发展方向(刘颖等,2020)。而针对大规模的图像数据,深度哈希检索(Ahmad等,2018)方法能够有效降低存储成本、提高查询速度,避免“维数灾难”。

目前,面向大规模的胸片图像检索仍然存在以

下问题:1)如何将深度哈希检索与海量胸片图像结合,解决“维数灾难”,提升检索效率。2)由于胸片图像类间相似度较高、病灶区域位置较小且难以分辨,导致特征提取时容易遗漏重要病灶信息,影响检索准确率。3)胸片图像噪声较大,仅考虑图像整体特征会由于包含病变区域之外的很大一部分噪声,导致对病变区域的误判;仅考虑图像局部特征会在获取病变信息分布全图时导致重要信息的丢失。4)针对现有的哈希方法学习的哈希码会丢失感兴趣区域和分类信息,导致大规模医学图像检索中小样本检索结果的排序质量降低。

为了解决这些问题,本文提出一种面向大规模胸片图像的深度哈希检索方法,流程图如图1所示。本文方法主要包括特征学习和哈希码优化两个模块。在特征学习模块中,除了获取图像的整体表示(即全局特征)外,构建空间注意模块获取胸片图像的病灶区域(即局部特征)。此外,将两种特征进行融合,相互补充,进一步完善图像的特征表示。在哈希码优化模块中,为了解决分类信息丢失以及排序问题,定义量化损失、二值交叉熵损失和正则化损失的联合函数进行优化学习。

本文的主要贡献如下:1)针对胸片图像类间相似度较高、病灶区域位置较小的问题,构建空间注意模块,挖掘胸片图像的显著区域表示,提取局部特征,增强哈希码的识别能力,提高后续检索的精度;

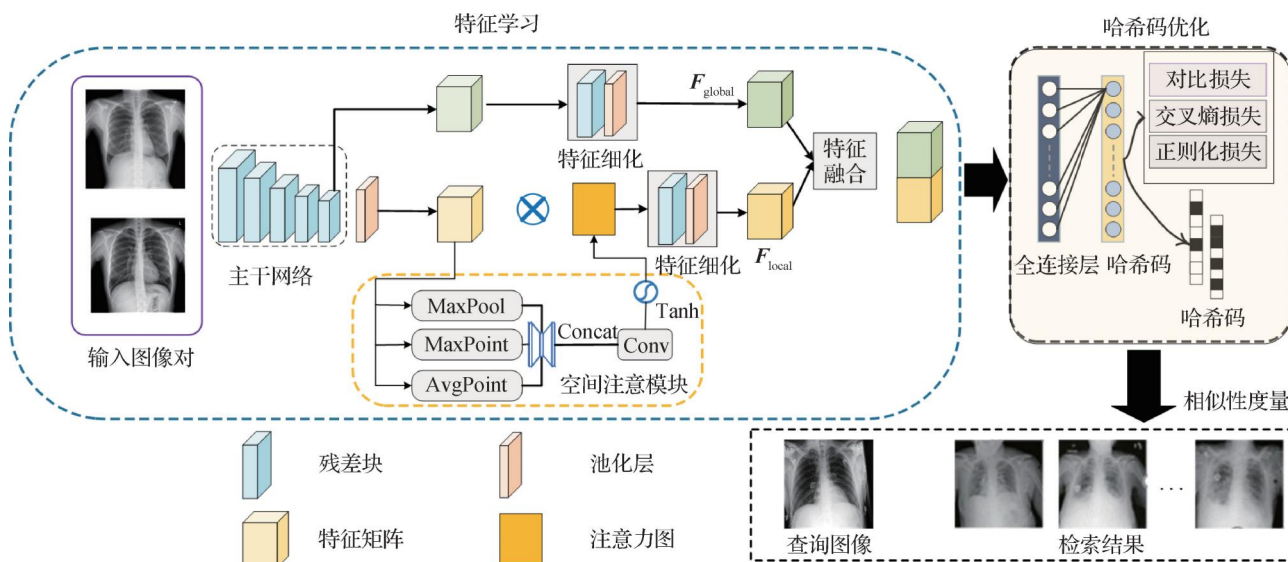


图1 大规模胸片图像深度哈希检索方法流程图

Fig. 1 The framework of deep hash retrieval methods for large scale chest X-ray images

2)在特征提取时,针对胸片图像噪声较大导致的信息丢失问题,定义特征融合模块,对细化后的局部特征和全局特征进行融合,解决病变信息丢失问题,以提高检索结果的准确率;3)在大规模胸片图像检索中,针对哈希学习方法因易丢失病灶区域和分类信息导致的结果排序质量较差问题,定义结合对比损失、交叉熵损失和正则化损失的联合函数,引入深度哈希训练网络中,使实值输出更接近2进制哈希码,提升了检索结果的排序质量。

1 相关工作

根据对数据的处理方式,哈希方法分为数据无关方法和数据依赖方法两类(刘颖等,2020)。数据无关方法通过独立训练数据得到哈希函数。其中,局部敏感哈希方法(Indyk和Motwani,1998)是常用的一种数据无关方法。数据依赖方法需要对训练数据进行学习以获得更好的检索性能,也称为学习哈希算法(Kong和Li,2012)。哈希深度学习方法在医学领域的应用取得了巨大成功。Liu等人(2014)采用锚图算法将乳房X光片特征压缩成紧凑的2进制码,以较高的概率保持图像特征空间的领域结构,从而减少存储空间和计算复杂度。Chen等人(2018)将多病症检索问题看做一个多标签哈希学习问题,提出一种基于顺序敏感的深度哈希方法,通过排序表和多标签分类获取多级语义相似度,结合排名损失和多标签分类损失更好地利用多标签信息。Jiang等人(2015)提出一种基于词汇树的图像检索框架,利用词汇树对特征信息进行细化算法,提高了X光片图像的检索精度。Shi等人(2018)给出基于成对的深度哈希排序算法,针对肺癌图像的特殊性设计成对向量,根据向量中的标签信息学习图像特征和2进制码,有效提高了分类准确性。Haq等人(2020)提出一种基于社区的大规模胸片图像检索方法,通过深度学习得到图编码,根据图编码将相似图像划分为社区,在社区中度量检索,在两个大胸片数据集上进行实验,得到了较好的检索性能。本文从提升胸片图像检索结果的排序质量角度入手,提出面向大规模胸片图像的深度学习哈希检索方法。

注意力机制已成功应用于卷积神经网络(convolutional neural network, CNN),显著提高了许多医学图像任务的性能(Nie等,2018),例如分割、分类和

检索等。Li等人(2019)提出一种基于注意力的青光眼检测网络,通过注意力预测子网将特征可视化局部病理区域,提高了青光眼检测及病理区域定位的性能。Woo等人(2018)提出一个注意模块CMBA(convolutional block attention module)。这个模块推导出的特征图包含通道和空间两个维度的信息,实验验证其优于仅使用通道注意力模型的方法。虽然性能得到了明显提升,但仍然会丢失一些小样本信息,导致结果排序质量不理想。Fang等人(2021)提出一个基于注意力的三元哈希(attention-based triplet hashing, ATH)网络以学习低维哈希码,该网络将注意力机制引入网络,结合了通道的最大值、平均值以及最大点值,联合上下文空间信息聚焦于感兴趣区域(regions of interest, ROI)区域。以上研究验证了空间注意机制可以通过捕获感兴趣区域信息来提高医学图像检索的性能。受ATH方法启发,本文在大规模胸片图像检索网络中设计了一个空间注意力机制,结合3个描述符聚焦胸片图像中的显著区域,获取局部特征。

对于基于内容的图像检索来说,图像特征提取是检索准确性的基石。由于单一特征只能从一个角度反映图像信息,因此在实际应用中,使用多特征融合比单一特征性能更为高效。全局特征捕捉图像的整体特征,而局部特征主要描述一组像素的特征,代表了细节。当病变信息只出现在图像相对较小的部分时,全局特征不能完全代表重要的视觉特征。Renita和Christopher(2020)提出使用grey wolf optimization-support vector machine进行医学图像检索,提取尺度和旋转不变特性以及纹理特征,之后采用单词包进行特征的映射。Qin等人(2019)提出一种全卷积稠密网络,分为编码和解码两个阶段,在编码部分提取图像丰富的语义特征,解码部分对带有语义特征的特征图进行上采样,并且使用联合距离将网络不同层的输出相加,作为最终图像检索的特征。Guan等人(2020)提出一种用于胸部疾病分类的卷积神经网络,利用注意力机制提取局部微小的病灶区域特征,再将局部和全局信息进行池化层融合来提高分类能力,得到更完整的一个图像表征。受以上工作启发,本文构建了特征融合模块,将提取并细化后的全局特征和局部特征进行融合,以获得更完整的图像表征,提高检索准确率。此外,由于模型训练过程中分类信息容易丢失,导致检索结果排序效

果不佳。本文在哈希码学习模块设计一个联合损失函数来学习哈希码, 用于保留分类信息, 使特征更易区分。

2 特征学习

由于残差学习的使用有助于缓解深度架构的梯度消失问题, 使训练更加稳定, 且其跳跃连接是无参数的, 不会增加模型的复杂性。本文采用 ResNet50 作为主干网络, 构建了空间注意模块提取输入胸片图像的全局—局部特征, 并引入特征融合及细化分支进行特征学习。

2.1 空间注意模块

为了关注胸部图像中的病灶区域特征, 提高哈希编码的特征表达能力, 本文构建了空间注意模块, 如图 1 所示。

首先, 使用 3 个描述符关注 ROI 区域。给定一个网络中间层特征图 $F \in \mathbf{R}^{W \times H \times K}$, 输出为注意力图 $M_F \in \mathbf{R}^{W \times H \times 1}$; 其中, K 、 H 和 W 分别表示特征图的通道数、高度和宽度。沿通道轴使用元素最大值和元素平均值操作, 生成两个不同的空间上下文信息 $F_{\max P}$ 和 F_{avgP} 。 $F_{\max P}$ 表示沿通道计算每个元素的最大值, F_{avgP} 表示沿通道计算每个元素的平均值。二者定义为

$$F_{\max P} = [f_1, \dots, f_i, \dots, f_{k \times k}], f_i = \max_{1 \leq c \leq k} \gamma_i(c) \quad (1)$$

$$F_{\text{avgP}} = [f_1, \dots, f_i, \dots, f_{k \times k}], f_i = \overline{\gamma_i(c)} \quad (2)$$

式中, $\gamma_i(c)$ 表示在 c 通道上 i 元素的响应值, f_i 表示在第 i 通道所取得的响应值。

其次, 为了关注信息部分, 引入特征 $F_{\max}$ 作为

补充, $F_{\max}$ 为特征经过最大池化层压缩处理得到。MaxPool 操作是对每个卷积层的最大局部响应进行编码, 而 MaxPoint 操作是针对特征图最大点的响应。将所得的 3 个描述符分别输入一个共享的多层感知机 (multilayer perception, MLP) 进行去噪, 然后将得到的输出连接并卷积成一个特征图, 再经过正切激活函数。具体操作为

$$M(F) = \omega(\gamma([MLP(F_{\max P}); \\ MLP(F_{\text{avgP}}); MLP(F_{\max})])) \quad (3)$$

式中, ω 定义为正切函数, γ 代表一个 3×3 的卷积操作。

在主干网络得到的特征图上, 空间注意力模块联合应用了 MaxPool、MaxPoint 和 AvgPoint 操作, 生成一个三通道特征图。对于每个像素, 计算 $F_{\max P}$ 补充了沿通道元素平均值的全局统计量, 并结合 F_{avgP} 更好地描述了通道的上下文特征。而最大池化操作 $F_{\max}$ 则避免了显著区域的边际值减弱, 提高了特征表达能力。因此, 构建的空间注意模块有助于获取胸片图像中的显著区域信息, 使网络在特征学习过程中专注于临床相关信息区域的特征提取。

2.2 特征融合

特征学习网络结构如图 2 所示。包含 7×7 的卷积层、 3×3 的最大池化层、残差块以及特征细化网络。其中, 选取的残差块来源于 ResNet 网络, 主要由一层 3×3 的卷积层、BatchNormal、ReLU (rectified linear unit) 组成, 如图 3 所示, 使用了跳跃连接, 解决梯度消失问题, 从而便于网络加深。特征细化网络为一个独立的网络结构, 由 1 个残差块和 1 个 3×3 平均池化层组成, 用于将关注的特征嵌入更高层的语义中。

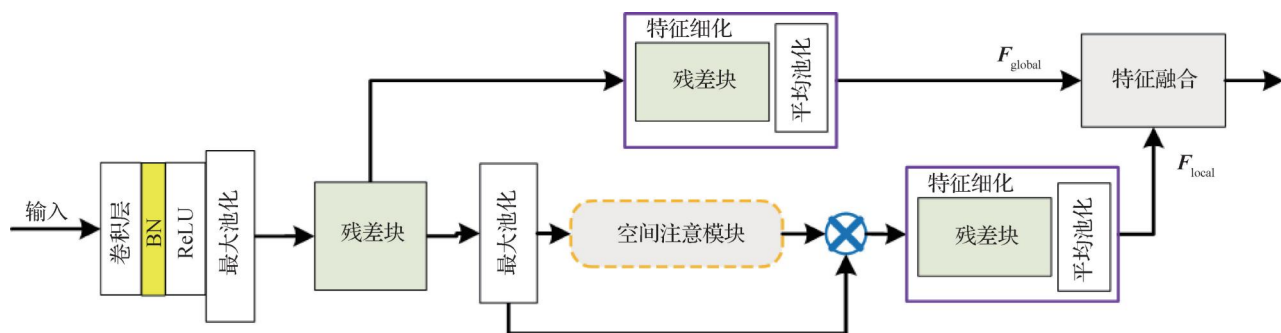


图 2 特征学习网络结构

Fig. 2 The structure of feature learning network

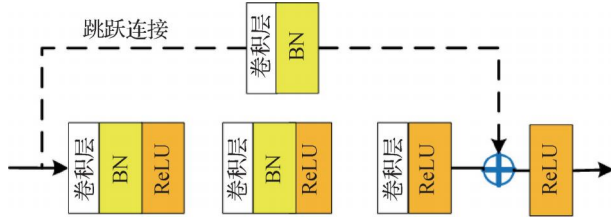


图3 残差块结构

Fig. 3 The structure of residual block

首先,提取全局特征。操作如下:将图像输入特征学习网络中,经过一系列处理后,在残差块处提取到一个整体深度特征 $F_i \in \mathbf{R}^{W \times H \times K}$ 。将该特征输入细化网络中进行全局特征的细化,生成特征 F_i^{global} , 具体为

$$F_i^{\text{global}} = f(F_i) \quad (4)$$

式中, $f(\cdot)$ 为特征细化操作。

其次,提取局部特征。局部特征主要关注于病变区域,有效解决了仅考虑图像整体特征时由于大量噪声导致的误判问题。将 F_i 输入一个大小为 3×3 、步长为 1 的最大池化层中,得到特征 F'_i 。通过最大池化处理,对 F_i 进行特征细节提取,突出影响较大特征的同时弱化一些不重要的特征,防止 ROI 区域的边际值减弱。然后,将 F'_i 输入空间注意模块(2.1 节)生成注意力图 $M_{F'_i} \in \mathbf{R}^{W \times H \times 1}$ 。将该注意力图 $M_{F'_i}$ 与 F'_i 相乘,得到注意局部特征 F''_i , 计算为

$$F''_i = M_{F'_i} \otimes F'_i \quad (5)$$

式中, $\otimes$ 表示每个通道上的哈达玛积。与全局特征一样,本文将提取的局部特征输入特征细化网络,进行局部特征细化,得到 F_i^{local} 特征,具体为

$$F_i^{\text{local}} = f(F''_i) \quad (6)$$

最后,进行全局和局部特征的融合。通过特征细化网络后,得到的 F_i^{global} 和 F_i^{local} 均为 256 维,将全局和局部特征按照维度进行拼接,然后将级联层连接到完全连接层(fully connected layers, FC),用于降维学习优化哈希码。

3 哈希码优化

为了得到高质量哈希码,改善结果排序质量,定义联合损失函数用于优化误差。常用的优化方法包括对比度量学习、三元组度量学习等(Li 等,2017)。对比度量学习可以更严格地对相似样本进行聚类,

使优化结果更为准确。因此,本文将提取的全局特征和局部特征进行连接输入到哈希码优化模块。哈希码优化模块包含 1 个完全连接层 FC 和 1 个哈希层,引入结合对比度量学习、哈希码正则化和交叉熵损失的联合损失函数进行优化。哈希层上的节点数设置为 k ,与所需哈希码的长度相等。对于 $i = 1, 2, \dots, k$ 的每一位哈希码,计算使用的函数为

$$b_i = \text{sign}(f) \quad (7)$$

式中, f 是卷积网络提取的高层语义特征。

为了缩小同类样本之间的距离,采用对比损失函数进行优化。该函数的输入为一对图像对,若为同一类样本,则通过优化缩小样本之间的距离;反之,存在一个阈值 m ,当不同样本之间距离小于 m 时,则互斥使样本之间的距离接近 m 。定义的对比损失函数为

$$L_c(f_1, f_2, y) = \frac{1}{2}(1 - y)D_w^2 + \frac{1}{2}y|\max(0, m - D_w)|^2 \quad (8)$$

式中, f_1 和 f_2 分别为输入样本对的高级语义特征; $\mathbf{1}$ 是一个数值全为 1 的向量;定义 n 维向量 y 表示输入样本对的监督信息,当样本对相似时, $y_i = 0$;否则, $y_i = 1$ 。 D_w 为 L2 归一化,用于度量哈希码之间的距离。具体地,当样本对相似时,调整参数最小化 f_1 与 f_2 之间的距离;当样本对不相似时,分两种情况,如果 f_1 与 f_2 之间的距离大于 m ,不进行优化;如果 f_1 与 f_2 之间的距离小于 m ,增大两者距离到 m ,其计算式为

$$D_w = \sqrt{\sum_{i=1}^n \sum_{j=1}^n (f_i - f_j)^2} \quad (9)$$

除了考虑对比损失,还定义了正则化项 L_r 使 2 进制码更接近于所需的哈希码。因此,为了优化映射函数,使网络输出近似于 2 进制哈希码,定义的目标函数表示为

$$L_r = \|\mathbf{f}_i\| - 1 \quad (10)$$

式中, f_i 是网络输出的高层语义特征。

为了生成保持语义的哈希码,在上述损失函数的基础上增加了一个交叉熵损失。该函数利用多标签信息,有效约束同类样本的特征分布距离,交叉熵损失函数定义为

$$L_{cb} = - \sum_{c=1}^C (y_{i,c} \ln \hat{y}_{i,c} + (1 - y_{i,c}) \ln(1 - \hat{y}_{i,c})) \quad (11)$$

式中, 真值 $y_{i,c} \in \{0, 1\}$ 表示样本 x_i 是否具有 c 类标签; $\hat{y}_{i,c}$ 为计算出属于 c 类标签的概率; C 是病例类别数目。通过累计相加每个类的交叉熵损失, 给出样本 x_i 的多标签分类损失。

综上所述, 本文定义的联合损失函数充分利用了胸片图像的标签信息和语义特征, 结合式(8)中的对比损失、式(10)的正则化损失和式(11)的交叉熵损失, 生成更具判别性的哈希码。本文哈希优化的联合损失函数表示为

$$L = L_c + \alpha L_r + \beta L_{ch} \quad (12)$$

式中, α 和 β 分别为正则化损失和交叉熵损失函数的权重。

在进行优化得到高质量哈希编码后, 通过相似性度量方法计算得到检索结果。本文采用汉明距离 (Hamming distance) 计算哈希码之间的相似度, 向量越相似, 汉明距离越小, 根据相似性对查询数据集进行升序排序, 返回排名前 k 的图像。汉明距离计算为

$$d(x, y) = \sum x[i] \oplus y[i] \quad (13)$$

式中, x 和 y 表示 K 维哈希编码, $i = 0, 1, \dots, k-1$, $d(x, y)$ 表示 x 与 y 之间的汉明距离, $\oplus$ 表示异或操作。至此, 通过以上方法输出最终的检索结果。

4 实验结果与分析

4.1 实验设置与评估标准

为了验证方法的性能, 使用两个公开可用的大规模胸片数据集进行实验。1) 美国国立卫生研究院 (National Institutes of Health, NIH) 的 ChestX-ray8 数据集 (Wang 等, 2017), 包含 30 805 名特殊患者的 112 120 幅正面 X-ray 图像, 每幅图像附带有相关的文本挖掘疾病标签。2) 美国斯坦福大学发布的 CheXpert 数据集 (Irvin 等, 2019), 包含 223 648 幅胸片和对应的病理报告。对检索任务, 本文将数据集中原始图像的 70% 用于训练, 30% 用于测试。

实验采用 Pytorch 实现提出的网络框架, 设备搭载 Intel Core i9-9900k CPU @ 3.60 GHz, NVIDIA GeForce RTX 2080Ti GPU 和 64 GB DDR4 2 666 MHz RAM。训练网络时, 使用带动量的小批量随机梯度下降 (stochastic gradient descent, SGD) 进行优化, 最小批次设置为 64, 初始学习率为 0.01, 权重衰减为

0.000 1, 动量为 0.9。共训练 50 个 epoch, 训练 20 个 epoch 后, 学习率下降到 0.001。通过交叉验证, 将正则化损失参数 α 和交叉熵损失参数 β 设置为 0.5。

评价指标采用哈希检索中常用的归一化折现累积增益 (normalized discounted cumulative gain, nDCG) 和平均精度 (mean average precision, mAP)。其中, nDCG 表示检索结果排名位置对检索结果的影响, mAP 指检索图像的平均准确率。

4.2 实验结果与性能分析

4.2.1 检索结果及对比分析

为了验证本文方法的检索性能, 与其他 6 种同任务的经典通用散列方法进行比较, 包括 SSH (semi-supervised hashing) (Wang 等, 2012)、ITQ (iterative quantization) (Gong 等, 2013)、DH (deep hashing) (Liong 等, 2015)、DSH (deep supervised hashing) (Liu 等, 2016)、DRH (deep residual hashing) (Conjeti 等, 2017) 和 ATH (Fang 等, 2021)。其中, DH、DSH、DRH 和 ATH 是基于深度学习的哈希方法, 其他是基于浅层的哈希方法。

为公平比较, 采用相同数据集进行实验。表 1 展示了本文方法与其他哈希方法在 ChestX-ray8 和 CheXpert 数据集上的 nDCG@100 和 mAP。可以看出, 基于深度学习的哈希方法比基于浅层的哈希方法表现更好, 本文方法的检索性能较其他深度学习的哈希方法具有一定优势。由于本文方法融合了全局和局部特征, 并构建了用于深度哈希检索的特征融合模块, 在 ChestX-ray 和 CheXpert 数据集上的 nDCG 比 ATH 方法分别提高了 4% 和 3%, mAP 值较 ATH 方法均提高了至少 5%。此外, 本文引入空间注意模块用于获取显著区域, 在两个数据集上检索返回 top-10 的图像中, mAP 值较 DSH 方法均有明显提升。实验结果表明, 空间注意模块能够聚焦局部病灶显著区域, 同时也验证了特征融合模块构建的必要性。

本文方法与同任务的哈希方法在 CheXpert 数据集上的特征计算时间、检索时间、训练时间和显存占用等方面的效率如表 2 所示, 表中数值是各项指标在整个测试集上的平均结果。可以看出, 传统哈希检索方法的训练时间明显快于深度哈希检索方法, 并且显存占用很小。而深度哈希检索方法都处于一个量级内。与 4 种深度哈希方法进行比较, 由

于嵌入了空间注意模块和特征融合模块,本文方法的网络训练时间和内存占用略高,特征计算时间

间和检索时间差距不大,但在结果准确率上有一定优势。

表 1 本文方法与同任务哈希方法在 ChestX-ray8 和 CheXpert 数据集上的性能比较

Table 1 Performance comparison between the hashing methods and our method with same task on ChestX-ray8 and CheXpert datasets

方法	ChestX-ray8 数据集		CheXpert 数据集	
	nDCG@ 100	mAP/%	nDCG@ 100	mAP/%
SSH(Wang 等,2012)	0. 13	0. 68	0. 15	0. 69
ITQ(Gong 等,2013)	0. 16	0. 65	0. 19	0. 71
DH(Liong 等,2015)	0. 12	0. 49	0. 13	0. 54
DSH(Liu 等,2016)	0. 18	0. 68	0. 2	0. 73
DRH(Conjeti 等,2017)	0. 19	0. 71	0. 22	0. 75
ATH(Fang 等,2021)	0. 21	0. 77	0. 24	0. 81
本文	0. 25	0. 83	0. 27	0. 86

表 2 本文方法与同任务哈希方法在 CheXpert 数据集上的效率比较

Table 2 Efficiency comparison between the hashing methods and our method with same task on CheXpert dataset

方法	特征计算时间/ms	检索时间/ms	训练时间/s	显存占用/MB
SSH(Wang 等,2012)	560	2. 8	4. 1	3 800
ITQ(Gong 等,2013)	670	2. 6	5	3 600
DH(Liong 等,2015)	64	4. 8	1 800	4 900
DSH(Liu 等,2016)	50	6	1 770	5 800
DRH(Conjeti 等,2017)	42	5	1 930	6 100
ATH(Fang 等,2021)	36	3. 4	2 300	6 700
本文	31	3	2 400	8 700

图 4 为本文方法部分查询图像的检索结果,每一行代表一组样本查询,第 1 列是查询图像,其他 5 列是前 5 个检索结果,每幅图像下面标注为疾病标签缩写,含义如表 3 所示。从图 4 可以看出,前两个检索示例展示了具有单个阳性标签的查询图像的检索结果,后两个查询示例展示了具有多个疾病标签的查询图像的检索结果。所有阳性疾病标签都可以从检索到的前几幅图像中检索出来,并能优先返回与查询图像共享更多疾病标签的图像。实验证明,本文方法能够提高检索结果的排序质量,同时返回高相似的结果。

4. 2. 2 可视化分析

本文提出的空间注意模块使用 3 个不同空间上下文信息描述特征图的有效性,即沿通道轴元素的

最大值 $F_{\max P}$ 、元素平均值 $F_{\text{avg}P}$ 和最大池化 $F_{\max}$ 。实验在 CheXpert 数据集进行,对比方法包括本文方法及两种变体:本文 w/o ($F_{\max P} + F_{\text{avg}P}$) 和本文 w/o $F_{\max}$ 。本文 w/o ($F_{\max P} + F_{\text{avg}P}$) 表示不考虑通道元素最大值和通道元素平均值;本文 w/o $F_{\max}$ 表示不考虑最大池化。实验结果如表 4 所示。由表 4 可以得出,不考虑 $F_{\max P}$ 和 $F_{\text{avg}P}$ 的情况下,本文方法的 mAP 值降低了 0. 18。不考虑 $F_{\max}$ 的情况下,mAP 降低了 0. 07。说明使用 3 个描述符能更有效地获取显著区域的信息。

为了说明空间注意模块在有效信息捕捉中的作用,图 5 给出了不同变体的热力图可视化结果。可以观察到,相比于不考虑 $F_{\max P}$ 和 $F_{\text{avg}P}$ 的情况,本文方法对于显著区域的关注更具体,减少了胸片的噪

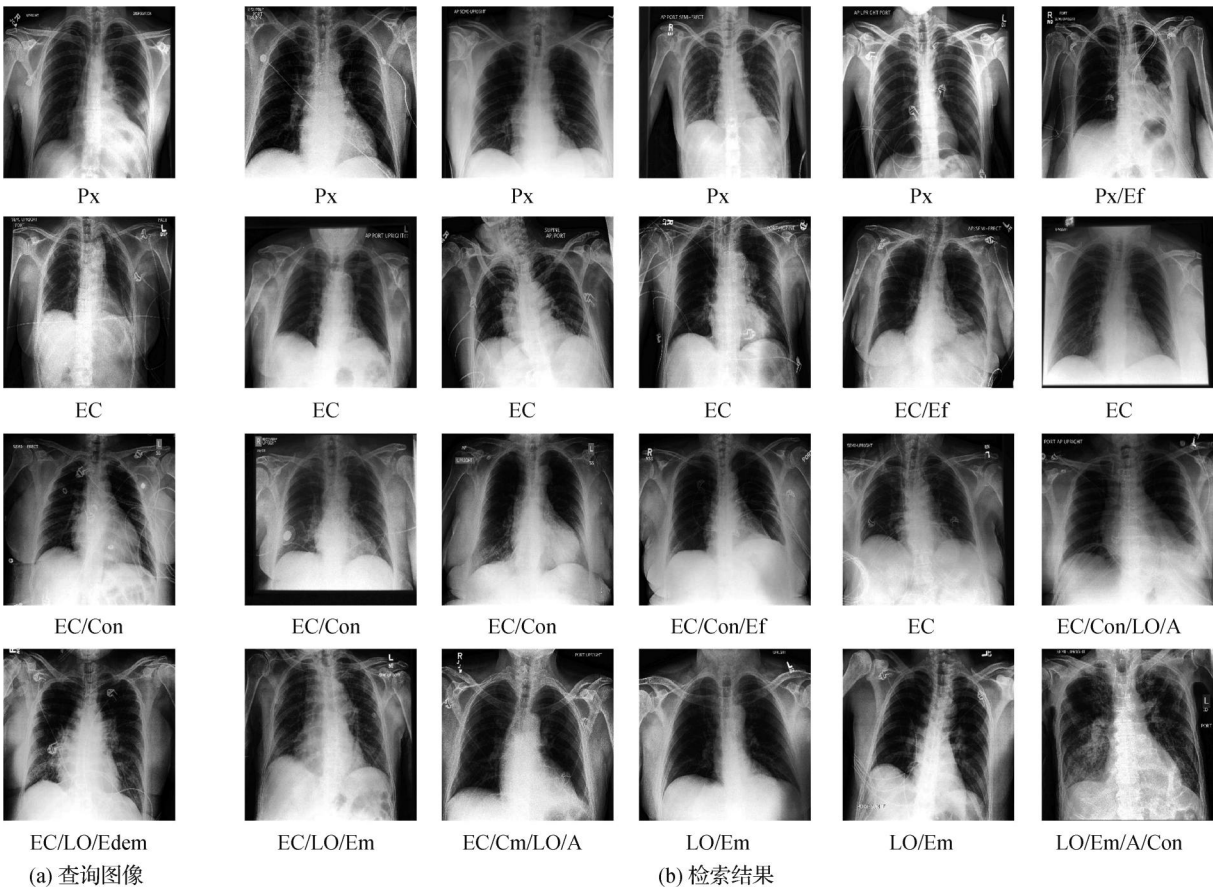


图 4 本文方法的检索结果

Fig. 4 Retrieval results of proposed method((a) query images;(b) retrieval results)

表 3 ChestX-ray8 和 CheXpert 数据集中疾病标签的含义
Table 3 The meaning of disease labels on ChestX-ray8 and CheXpert datasets

疾病标签	ChestX-ray8 数据集		CheXpert 数据集	
	缩写	中文	缩写	中文
atelectasis	A	肺不张	A	肺不张
enlarged cardiomeastinum	EC	扩张型心肌病	-	-
cardiomegaly	Cm	心肥大	Cm	心肥大
pneumothorax	Px	气胸	Px	气胸
pleural effusion	Ef	胸腔积液	Ef	胸腔积液
consolidation	Con	实变	Con	实变
edema	Em	水肿	Em	水肿
lung opacity	LO	肺不透明	LO	肺不透明
infiltration	-	-	Infi	浸润
mass	-	-	Mass	肿块
nodule	-	-	Nodule	结节
pneumonia	-	-	Pne2	肺炎

注：“-”表示数据集未涉及。

表 4 不同空间上下文描述特征图的有效性对比
Table 4 Effectiveness comparison of different spatial contexts for describing feature maps

方法	mAP/%
本文	0.83
本文 w/o ($F_{\max P} + F_{\text{avgP}}$)	0.65
本文 w/o $F_{\max}$	0.76

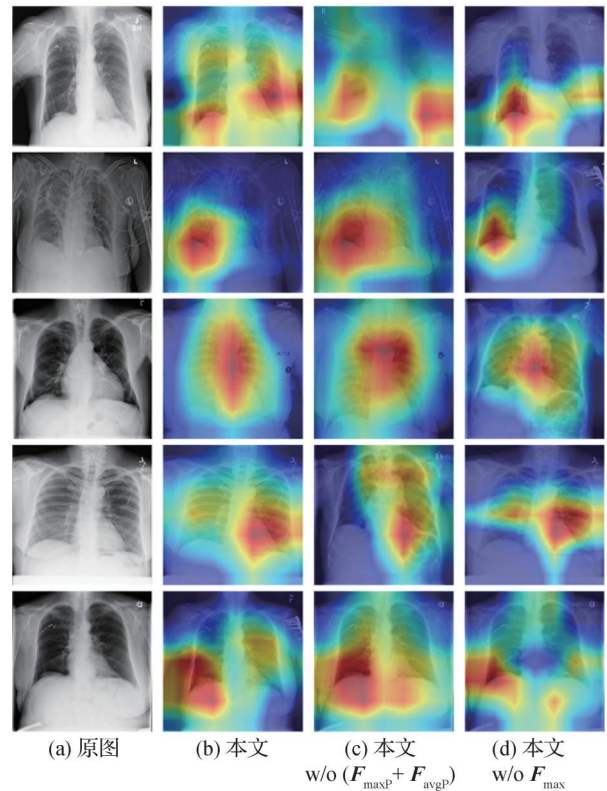


图 5 不同方法的热图结果

Fig. 5 Heat maps for different methods

((a) original images; (b) our method; (c) our method w/o ($F_{\max P} + F_{\text{avgP}}$); (d) our method w/o $F_{\max}$)

声影响;相比于不考虑 $F_{\max}$ 的情况,本文方法对显著区域的关注更全面,有效解决了病灶的遗漏问题。

为进一步验证空间注意模块的有效性,在 CheXpert 数据集上进行实验,对疾病定位进行性能比较,定位结果如图 6 所示。对 Nodule 和 Mass 等疾病,病变区域定位的准确性优于另外两种方法。对 Cm 疾病,病变区域的定位结果与其他方法差异不大,主要原因在于该方法主要侧重小的病变区域。实验结果表明,本文引入的空间注意模块对小病变区域的定位具有一定优势,能够有效捕捉病变区域,增强哈希码的可识别性。

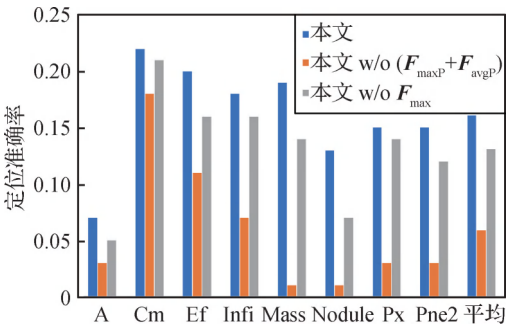


图 6 疾病定位结果

Fig. 6 The results of disease location

4. 2. 3 消融实验分析

为验证融合全局特征与局部特征的有效性,采用 mAP 和 nDCG@100 作为评价标准,实验结果如表 5 所示。从表 5 可以看出,本文方法融合特征的 mAP 值相比单一特征至少提高了 4%,检索结果最好。nDCG@100 为 0.27,也超过单一特征的性能,获得了更好的排序结果。实验结果说明,通过全局特征和局部特征的融合,不仅关注整体特征,还考虑局部病变区域,使网络中的特征包含了小样本以及同类的相关信息,能够缩小类内差距,有益于提高检索的准确率。

表 5 使用不同特征的性能分析
Table 5 Performance analysis of using different features

特征	mAP/%	nDCG@100
局部特征	0.77	0.12
全局特征	0.79	0.17
融合全局特征与局部特征(本文)	0.83	0.27

训练损失的定性结果如图 7 所示。从图 7 可以看出,网络收敛速度开始较快,但慢慢振荡,最终趋于稳定拟合。

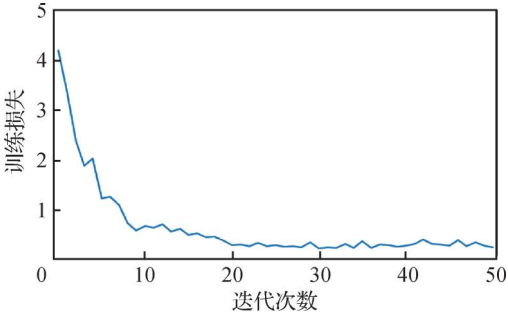


图 7 训练损失结果

Fig. 7 Qualitative results of training loss

5 结 论

为解决大规模胸片图像的检索难题, 提出一种面向大规模胸片图像的深度哈希检索方法。首先在网络中构建空间注意模块, 提取局部特征, 增强哈希码的识别能力。然后针对病灶信息丢失问题, 定义特征融合模块, 将细化后的全局特征和局部特征进行融合。最后结合 3 个损失函数进行哈希码优化, 提升检索结果准确率。实验结果显示, 在 ChestXay-8 和 CheXpert 数据集上, 检索准确率分别提高了 6% 和 5%, 表明本文方法在胸片图像检索中具有明显优势。

虽然本文方法能够对医学图像进行精确检索, 但是仍然存在需要改进的地方。一方面, 可以参照现有的哈希方法, 对损失函数进一步优化, 使其更易于区分小样本图像; 另一方面, 为了关注 ROI 区域, 可以尝试调整网络的组合顺序。在未来工作中, 将围绕这些方面进行研究, 致力于在胸片图像检索中获得更好的性能。

参考文献 (References)

- Ahmad J, Muhammad K and Baik S W. 2018. Medical image retrieval with compact binary codes generated in frequency domain using highly reactive convolutional features. *Journal of Medical Systems*, 42(2): #24 [DOI: 10.1007/s10916-017-0875-4]
- Chen Z X, Cai R J, Lu J W, Feng J J and Zhou J. 2018. Order-sensitive deep hashing for multimorbidity medical image retrieval//Proceedings of the 21st International Conference on Medical Image Computing and Computer Assisted Intervention. Granada, Spain; Springer: 620-628 [DOI: 10.1007/978-3-030-00928-1_70]
- Conjeti S, Roy A G, Katouzian A and Navab N. 2017. Hashing with residual networks for image retrieval//Proceedings of the 20th International Conference on Medical Image Computing and Computer Assisted Intervention. Quebec City, Canada; Springer: 541-549 [DOI: 10.1007/978-3-319-66179-7_62]
- Fang J S, Fu H Z and Liu J. 2021. Deep triplet hashing network for case-based medical image retrieval. *Medical Image Analysis*, 69: #101981 [DOI: 10.1016/j.media.2021.101981]
- Gong Y C, Lazebnik S, Gordo A and Perronnin F. 2013. Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12): 2916-2929 [DOI: 10.1109/TPAMI.2012.193]
- Guan Q J, Huang Y P, Zhong Z, Zheng Z D, Zheng L and Yang Y. 2020. Thorax disease classification with attention guided convolutional neural network. *Pattern Recognition Letters*, 131: 38-45 [DOI: 10.1016/j.patrec.2019.11.040]
- Haq N F, Moradi M and Wang Z J. 2020. A deep community based approach for large scale content based X-ray image retrieval. *Medical Image Analysis*, 68(2): #101847 [DOI: 10.1016/j.media.2020.101847]
- Hwang K H, Lee H and Choi D. 2012. Medical image retrieval: past and present. *Healthcare Informatics Research*, 18(1): 3-9 [DOI: 10.4258/hir.2012.18.1.3]
- Indyk P and Motwani R. 1998. Approximate nearest neighbors: towards removing the curse of dimensionality//Proceedings of the 30th Annual ACM Symposium on Theory of Computing. Dallas, USA; Association for Computing Machinery: 604-613 [DOI: 10.1145/276698.276876]
- Irvin J, Rajpurkar P, Ko M, Yu Y F, Ciurea-Illcus S, Chute C, Marklund H, Haghighi B, Ball R, Shpankaya K, Seekins J, Mong D A, Halabi S S, Sandberg J K, Jones R, Larson D B, Langlotz C P, Patel B N, Lungren M P and Ng A Y. 2019. CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1): 590-597 [DOI: 10.1609/aaai.v33i01.3301590]
- Jiang M L, Zhang S T, Li H S and Metaxas D N. 2015. Computer-aided diagnosis of mammographic masses using scalable image retrieval. *IEEE Transactions on Biomedical Engineering*, 62(2): 783-792 [DOI: 10.1109/tbme.2014.2365494]
- Kashif M, Raja G and Shaukat F. 2020. An efficient content-based image retrieval system for the diagnosis of lung diseases. *Journal of Digital Imaging*, 33(4): 971-987 [DOI: 10.1007/s10278-020-00338-w]
- Kong W H and Li W J. 2012. Isotropic hashing//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA; Curran Associates Inc.: 1646-1654
- Li L, Xu M, Wang X F, Jiang L and Liu H R. 2019. Attention based glaucoma detection: a large-scale database and CNN model//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA; IEEE: 10563-10572 [DOI: 10.1109/CVPR.2019.01082]
- Li Q, Sun Z N, He R and Tan T N. 2017. Deep supervised discrete hashing//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA; Curran Associates Inc.: 2479-2488
- Li Z Y, Zhang X F, Müller H and Zhang S T. 2018. Large-scale retrieval for medical image analytics: a comprehensive review. *Medical Image Analysis*, 43: 66-84 [DOI: 10.1016/j.media.2017.09.007]
- Liong V E, Lu J W, Wang G, Moulin P and Zhou J. 2015. Deep hashing for compact binary codes learning//Proceedings of 2015 Confer-

- ence on Computer Vision and Pattern Recognition (CVPR). Boston, USA; IEEE: 2475-2483 [DOI: 0. 1109/CVPR. 2015. 7298862]
- Liu H M, Wang R P, Shan S G and Chen X L. 2016. Deep supervised hashing for fast image retrieval//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA; IEEE: 2064-2072 [DOI: 10. 1109/CVPR. 2016. 227]
- Liu J J, Zhang S T, Liu W, Zhang X F and Metaxas D N. 2014. Scalable mammogram retrieval using Anchor Graph Hashing//Proceedings of the 11th IEEE International Symposium on Biomedical Imaging (ISBI). Beijing, China; IEEE: 898-901 [DOI: 10. 1109/ISBI. 2014. 6868016]
- Liu Y, Cheng M, Wang F P, Li D X, Liu W and Fan J L. 2020. Deep Hashing image retrieval methods. Chinese Journal of Image and Graphics, 25(7): 1296-1317 (刘颖, 程美, 王富平, 李大湘, 刘伟, 范九伦. 2020. 深度哈希图像检索方法综述. 中国图形图象学报, 25(7): 1296-1317) [DOI: 10. 11834/jig. 190518]
- Nie D, Gao Y Z, Wang L and Shen D G. 2018. ASDNet: attention based semi-supervised deep networks for medical image segmentation//Proceedings of the 21st International Conference on Medical Image Computing and Computer Assisted Intervention. Granada, Spain; Springer: 370-378 [DOI: 10. 1007/978-3-030-00937-3_43]
- Qin P L, Li Q, Zeng J C, Liu H Y and Cui Y H. 2019. Fully convolutional-based dense network for lung nodule image retrieval algorithm. International Journal of Performability Engineering, 15(1): 326-336 [DOI: 10. 23940/ijpe. 19. 01. p33. 326336]
- Renita D B and Christopher C S. 2020. Novel real time content based medical image retrieval scheme with GWO-SVM. Multimedia Tools and Applications, 79(23): 17227-17243 [DOI: 10. 1007/s11042-019-07777-w]
- Shi X S, Sapkota M, Xing F Y, Liu F J, Cui L and Yang L. 2018. Pairwise based deep ranking hashing for histopathology image classification and retrieval. Pattern Recognition, 81: 14-22 [DOI: 10. 1016/j. patcog. 2018. 03. 015]
- Silva W, Poellinger A, Cardoso J S and Reyes M. 2020. Interpretability-guided content-based medical image retrieval//Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention. Lima, Peru; Springer: 305-314 [DOI: 10. 1007/978-3-030-59710-8_30]
- Wang J, Kumar S and Chang S F. 2012. Semi-supervised hashing for large-scale search. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(12): 2393-2406 [DOI: 10. 1109/TPAMI. 2012. 48]
- Wang X S, Peng Y F, Lu L, Lu Z Y, Bagheri M and Summers R M. 2017. ChestX-Ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE: 3462-3471 [DOI: 10. 1109/cvpr. 2017. 369]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany; Springer: 3-19 [DOI: 10. 1007/978-3-030-01234-2_1]
- Wu M L, Chen Q and Sun Q S. 2013. Medical image retrieval by graph-based semi-supervised learning. Journal of Computer-Aided Design and Computer Graphics, 25(9): 1354-1360 (吴梦麟, 陈强, 孙权森. 2013. 基于图半监督学习的医学图像检索. 计算机辅助设计与图形学学报, 25(9): 1354-1360) [DOI: 10. 3969/j. issn. 1003-9775. 2013. 09. 012]

作者简介

管安娜,女,硕士研究生,主要研究方向为医学图像处理和计算机视觉。E-mail: 736116785@qq.com

刘骊,通信作者,女,教授,博士生导师,主要研究方向为计算机辅助设计与图形学、图像处理和计算机视觉。

E-mail: kmust_mary@163.com

付晓东,男,教授,博士生导师,主要研究方向为服务计算、决策理论与方法。E-mail: xiaodong_fu@hotmail.com

刘利军,男,副教授,硕士生导师,主要研究方向为医学图像处理和信息检索。E-mail: cloneiq@126.com

黄青松,男,教授,硕士生导师,主要研究方向为机器学习和医学智能信息系统。E-mail: 1912443688@qq.com