

文章编号:1671-6833(2025)02-0060-07

# 基于交叉量化和样本校正的自监督遥感图像检索

潘丽丽<sup>1</sup>, 瞿栋梁<sup>2</sup>, 尹晶晶<sup>2</sup>, 马雪强<sup>1</sup>

(1. 中南林业科技大学 计算机与数学学院, 湖南 长沙 410000; 2. 中南林业科技大学 电子信息与物理学院, 湖南 长沙 410000)

**摘要:** 自监督遥感图像检索模型由于标签缺失导致使用错误的样本对进行学习, 从而产生抽样偏差, 影响图像表征的准确性, 为此提出一种基于交叉量化和样本校正的自监督遥感图像检索模型(CQSC)。首先, 为了降低数据存储和处理负载, 将传统对比学习中的映射层与乘积量化相结合, 压缩高维图像数据, 提高了检索效率; 其次, 使用交叉学习策略, 最大化检索模型中特征映射前后的交叉相似性, 增强模型的特征生成能力和检索精度; 最后, 针对自监督遥感图像检索中因标签缺失导致的抽样偏差, 设计自适应纠错标签, 标注训练样本, 校正训练过程中存在的错误负样本。在 UC Merced 和 EuroSAT 数据集上进行了大量实验, 结果表明: 与 PLSH 方法相比, 在 UC Merced 数据集上, 所提方法  $mAP@20$  平均提升了 2.52 个百分点; 在 EuroSAT 数据集上取 64 bits 时, 所提方法  $mAP@100$  提升了 3.83 个百分点。

**关键词:** 遥感图像检索; 对比学习; 乘积量化; 交叉学习

**中图分类号:** P407.8; TP183

**文献标志码:** A

**doi:** 10.13705/j.issn.1671-6833.2024.06.007

随着机器学习<sup>[1]</sup>的兴起, 基于卷积神经网络的遥感图像检索方法<sup>[2]</sup>被提出。由于标签制作费时费力, 研究者们提出无监督遥感图像检索<sup>[3]</sup>, 但这又引发了另一个问题: 标签的缺失会导致从训练数据中随机抽取的样本存在负样本与正样本相似的情况<sup>[4]</sup>。Chuang 等<sup>[5]</sup>将这种现象称为抽样偏差, 并且用实验证明抽样偏差会影响检索模型的最终性能。

另一方面, 传统遥感图像检索直接将特征映射到高维欧氏空间, 导致图像特征变成上千维的实值向量, 这不仅增加了存储成本, 还降低了检索速度。

针对这个问题, 本文提出一种基于交叉量化和样本校正的自监督遥感图像检索模型(简称 CQSC)。首先, 使用乘积量化替换对比学习中的映射层, 从而产生比哈希更丰富的距离表示, 有效降低量化误差对模型性能的影响, 在保证模型的准确性和鲁棒性的同时提升检索速度。其次, 引入交叉学习策略, 最大化对比学习中映射层前后图像描述符之间的交叉相似性, 使模型提取的图像特征更具有判别力, 帮助 CQSC 进一步提高检索精度。最后, 通

过计算正样本与负样本之间的矢量距离, 使用自适应纠错标签标注出训练样本中的错误负样本, 避免模型将正样本误认为负样本, 以此来提高模型的检索精度。

## 1 相关工作

### 1.1 深度哈希图像检索

近似最近邻(approximate nearest neighbor, ANN)搜索<sup>[6]</sup>以低存储成本和快速搜索等优势在遥感图像检索研究中备受关注。ANN 研究包含两种主流方法: 哈希(Hashing)<sup>[7]</sup>和矢量量化(vector quantization, VQ)<sup>[8]</sup>。二者虽然都是将高维图像数据转换为紧凑的二进制代码, 但是哈希只是使用汉明距离<sup>[9]</sup>计算二进制代码之间的距离, 无法表示复杂的距离关系。乘积量化(product quantization, PQ)可以快速准确地提供检索结果。Liang 等<sup>[10]</sup>提出了一种基于 PQ 的深度哈希框架, 利用基于 Soft-max 的可微分 PQ 分支来吸收与该类相关的预先建立的 PQ 代码, 不需要复杂的矩阵操作, 就能够熟练

收稿日期: 2024-10-06; 修订日期: 2024-12-02

基金项目: 湖南省自然科学基金资助项目(2021JJ31164); 湖南省教育厅科学研究重点项目(22A0195)

作者简介: 潘丽丽(1977—), 女, 湖南长沙人, 中南林业科技大学教授, 博士, 主要从事机器学习和图像处理研究, E-mail: lily\_pan@163.com。

引用本文: 潘丽丽, 瞿栋梁, 尹晶晶, 等. 基于交叉量化和样本校正的自监督遥感图像检索[J]. 郑州大学学报(工学版), 2025, 46(2): 60-66. (PAN L L, QU D L, YIN J J, et al. Self-supervised remote sensing image retrieval based on cross-quantization and sample correction[J]. Journal of Zhengzhou University (Engineering Science), 2025, 46(2): 60-66.)

地培养高度独特和紧凑的哈希码。

1.2 自监督遥感图像检索

在自监督遥感图像检索<sup>[11]</sup>中,计算机使用遥感图像数据自动学习特征表示。与传统的监督学习方法相比,自监督遥感图像检索不需要标注数据,可以从未标注数据中学习特征表示,因此具有更强的泛化能力和更高的效率。对比学习<sup>[12]</sup>是一种广泛应用于自监督学习中的方法,但是由于抽样偏差的存在,导致训练使用的负样本与目标样本相似,从而导致检索精度降低。Chuang 等<sup>[5]</sup>提出一种近似理论无偏的对比损失,并没有直接针对抽样偏差问题提出解决方案。NNCLR<sup>[13]</sup>对数据特征进行排序来缓解抽样偏差引起的性能下降,但是依赖于预定义的样本数量。

2 基于交叉量化和样本校正的自监督遥感图像检索

如图 1 所示,遥感图像经过数据增强可以得到不同的增强版本,所用特征增强包含随机裁剪、随机旋转、高斯模糊等。通过将特征增强后的样本映射进同一纬度得到其特征向量。遥感图像  $X_i$  在经过随机数据增强之后,得到图像  $X_i$  的不同增强版本  $\{X_{i1}, X_{i2}\}$ ,使用 ResNet50 将特征增强后的样本  $\{X_{i1}, X_{i2}\}$  映射到同一纬度,从而得到遥感图像的特征向量  $\{x_{i1}, x_{i2}\}$ ,  $x_{i1} \in \mathbf{R}^D, x_{i2} \in \mathbf{R}^D$ ,  $D$  为特征向量的原始维度。

关于遥感图像检索的量化,乘积量化通过笛卡

尔积运算将码字规范成一个码本,将特征向量  $x_{i1}$  分成  $m$  个独立的子特征向量,即  $x_{i1} = \{x_{i11}, x_{i12}, \dots, x_{i1m}\}$ ,  $x_{i1m} \in \mathbf{R}^{D/m}$ 。量化头  $G$  由  $m$  个独立的子量化器  $g_i$  组成,  $G = \{g_1, g_2, \dots, g_m\}$ 。  $g_i$  又由  $K$  个不同的码字组成,  $g_i = \{g_{i1}, g_{i2}, \dots, g_{iK}\}$ ,  $g_{iK} \in \mathbf{R}^{D/m}$ 。通过量化头  $G$  分别对这  $m$  个子特征向量进行量化索引。将遥感图像特征向量  $x_{i1}$  分割后的  $m$  个子特征向量,在对应的子量化器中找到最接近的码字并替换,将替换的码字拼接成一个新的向量,得到特征向量  $x_{i1}$  自身的量化描述符  $\eta_{i1}$ 。对于增强样本  $X_{i2}$  进行同样的处理可以得到其量化描述符  $\eta_{i2}$ 。

通过计算不同样本量化描述符之间的距离,将距离转化成相似概率,加权得到最终的概率分布。通过样本特征空间的相对分布赋予样本纠错标签,标注训练样本中的错误负样本。

交叉量化通过压缩图像数据提供高效的特征表示,自适应纠错标签通过概率分布调整优化模型处理错误负样本的能力。二者结合显著提高了自监督遥感图像检索的性能,实现了检索效率和准确度的双重提升。

2.1 基于交叉量化的遥感图检索

基于交叉量化的自监督遥感图像检索任务的主要算法过程如下。

首先,计算子特征向量与对应的子量化器中码字之间的距离,例如特征向量  $x_{i1}$  的第  $n$  个子特征向量  $x_{i1n}$  与第  $n$  个子量化器中码字的距离为

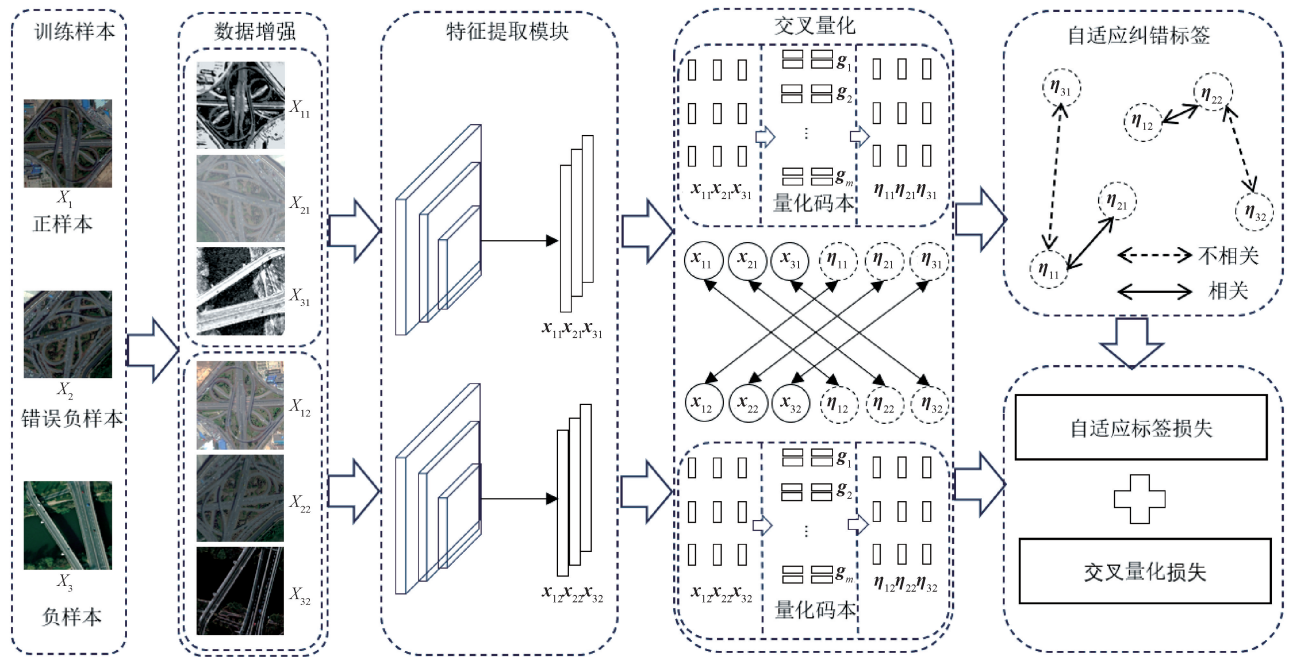


图 1 基于交叉量化和样本校正的自监督遥感图像检索框架

Figure 1 Self-supervised remote sensing image retrieval framework based on cross-quantization and sample correction

$$d_{nk} = \| \mathbf{x}_{i1n} - \mathbf{g}_{nk} \|_2^2, k \in \{1, 2, \dots, K\}. \quad (1)$$

式中:  $\mathbf{g}_{nk}$  为子量化器  $\mathbf{g}_n$  中的第  $k$  个码字;  $\| \cdot \|_2^2$  为  $L_2$  范数。

通过计算每个切分后的  $\mathbf{x}_{i1n}$  和  $\mathbf{g}_{nk}$  之间的欧氏距离,使用 Softmax 函数将距离转换为相应的概率值,如式(2)所示:

$$\boldsymbol{\eta}_{i1n} = \operatorname{argmax}(\mathbf{d}_{nk}, \mathbf{g}_n). \quad (2)$$

式中:  $\boldsymbol{\eta}_{i1n}$  为属于第  $n$  个子量化器  $\mathbf{g}_n$  中最接近子特征向量  $\mathbf{x}_{i1n}$  的码字。

其次,将  $\boldsymbol{\eta}_{i1n}$  看作子量化器  $\mathbf{g}_n$  中码字的指数加权和,将子量化器中的码字近似量化输出,最接近子特征向量的码字将会被输出作为子特征向量的量化表示,如式(3)所示:

$$\boldsymbol{\eta}_{i1n} = \left( \frac{\sum_{k=1}^K \frac{\exp \frac{-\| \mathbf{x}_{i1n} - \mathbf{g}_{nk} \|_2^2}{\phi_g}}{\sum_{k'=1}^K \exp \frac{-\| \mathbf{x}_{i1n} - \mathbf{g}_{nk'} \|_2^2}{\phi_g}}}{\sum_{k'=1}^K \exp \frac{-\| \mathbf{x}_{i1n} - \mathbf{g}_{nk'} \|_2^2}{\phi_g}} \right) \cdot \mathbf{g}_{nk} \quad (3)$$

式中:  $\phi_g$  为非负温度参数,设置为 5.0;  $\| \cdot \|_2^2$  为计算每个切分后的子特征向量  $\mathbf{x}_{i1n}$  和子量化器的码字  $\mathbf{g}_{nk}$  之间的余弦距离。

最后,将得到的码字拼接在一起,就是特征向量  $\mathbf{x}_{i1}$  乘积量化后的量化描述符  $\boldsymbol{\eta}_{i1}$ 。

何悦等<sup>[14]</sup>提出一种基于深度多相似性哈希(DMSH)的遥感图像检索方法,通过挖掘图像间的多种相似信息生成哈希码。吴刚等<sup>[15]</sup>提出一种基于级联池化自注意力模块的遥感影像检索方法,采用监督核哈希生成哈希码。相较于这些基于传统哈希的遥感图像检索,本文提出的交叉量化降低量化阶段的信息损失,减少距离计算的时间复杂度,更好地保留原始数据的结构和相似性。

为了更好地学习特征向量和码字,增加码字的泛化能力,引入交叉学习策略。对于同一幅遥感图像来说,使用不同的随机增强得到的特征向量和量化描述符之间是相关联的,与批量样本内的其他图像是不相关的。基于这一思想,设计了一个交叉量化对比损失函数来学习相关的实例对,如式(4)所示:

$$L_{\text{eq}} = \frac{1}{2N} \sum_{n=1}^N [L_c(\mathbf{x}_{i1}, \boldsymbol{\eta}_{i2}) + L_c(\mathbf{x}_{i2}, \boldsymbol{\eta}_{i1})]; \quad (4)$$

$$L_c(\mathbf{x}_{i1}, \boldsymbol{\eta}_{i2}) = -\log \frac{\exp \frac{S(\mathbf{x}_{i1}, \boldsymbol{\eta}_{i2})}{\phi_{\text{eq}}}}{\sum_{n=1, n \neq j}^N \exp \frac{S(\mathbf{x}_{i1}, \boldsymbol{\eta}_{in})}{\phi_{\text{eq}}}}. \quad (5)$$

式中:  $S$  表示余弦相似性;  $\phi_{\text{eq}}$  为交叉量化损失函数的非负温度参数。

交叉量化利用不同随机增强的遥感图像特征,通过乘积量化促进特征准确量化,增强了模型对遥感图像中复杂地物和场景的区分能力,优化了检索结果的相关性和准确性。

## 2.2 自适应纠错标签

如上所述,对比学习在训练过程中会主观地认定每个样本都是一个独立的个体,互相没有关联,同一批次中只有一个正样本,其余全是负样本,从而导致抽样偏差。针对这一问题,设计一个纠错标签。具体步骤如下。

**步骤 1** 对于遥感图像的特征向量  $\mathbf{x}_{i1}$ , 计算特征向量  $\mathbf{x}_{i1}$  的量化描述符  $\boldsymbol{\eta}_{i1}$  与批量样本  $\{\mathbf{x}_{11}, \mathbf{x}_{21}, \dots, \mathbf{x}_{N1}\}$  的量化描述符  $\{\boldsymbol{\eta}_{11}, \boldsymbol{\eta}_{21}, \dots, \boldsymbol{\eta}_{N1}\}$  之间的余弦相似度  $S_{ij}$ :

$$S_{ij} = \frac{\boldsymbol{\eta}_{i1}^T \boldsymbol{\eta}_{j1}}{\| \boldsymbol{\eta}_{i1} \|_2^2 \| \boldsymbol{\eta}_{j1} \|_2^2}, j = 1, 2, \dots, N. \quad (6)$$

式中:  $\boldsymbol{\eta}_{i1}$  和  $\boldsymbol{\eta}_{j1}$  为遥感图像的量化描述符;  $\| \cdot \|_2^2$  为  $L_2$  范数;  $\boldsymbol{\eta}_{i1}^T$  为  $\boldsymbol{\eta}_{i1}$  的转置。

**步骤 2** 通过余弦相似度  $S_{ij}$  来构建批量样本的相似分布  $Q_i$ , 如式(7)所示:

$$Q_i = \frac{\exp \frac{S_{ij}}{\phi_a}}{\sum_j \exp \frac{S_{ij}}{\phi_a}}. \quad (7)$$

式中:  $\phi_a$  为控制样本相似分布的参数。

**步骤 3** 将置信度定义为相对分布  $Q_i$  的归一化熵  $E$ , 如式(8)所示:

$$E = 1 - \frac{H(Q_i)}{\log(N-1)}. \quad (8)$$

式中:  $H(Q_i)$  用于计算相对分布  $Q_i$  的香农熵;  $N$  为样本数量。

**步骤 4** 使用相对分布  $Q_i$  本身作为样本的纠错标签,并使用相对分布  $Q_i$  的归一化熵  $E$  对自适应纠错标签进行加权,纠错标签  $y$  定义为

$$y_i = \begin{cases} 1, i = j; \\ E \cdot Q_i, i \neq j. \end{cases} \quad (9)$$

最终批量样本的纠错标签损失  $L_a$  为

$$L_a = \frac{1}{N} \sum_{i=1}^N \left( 2 - 2 \sum_{j=1}^N \frac{\boldsymbol{\eta}_{i1}^T \boldsymbol{\eta}_{j1}}{\| \boldsymbol{\eta}_{i1} \|_2^2 \| \boldsymbol{\eta}_{j1} \|_2^2} \cdot y_i \right). \quad (10)$$

Chuang 等<sup>[5]</sup>提出一种近似理论无偏的对比损失,并没有直接针对抽样偏差问题提出解决方案。NNCLR<sup>[13]</sup>对数据特征使用排序来缓解抽样偏差引起的性能下降,但是依赖于预定义的样本数量。这些方法并不适用于所有的遥感图像类型和场景,本



文通过引入纠错标签,不需要引入预定义数据,针对抽样偏差问题提供了解决方案,通过动态调整标签来减少这种错误,从而提高检索的准确性。

### 2.3 多损失函数学习

在训练学习阶段,本文采用基于交叉量化的对比损失函数和自适应纠错标签损失来共同更新模型参数。基于交叉量化的损失函数学习遥感图像的深度特征和码字来区分图像内容,保证遥感图像的深度特征可以保留在量化编码中,如式(4)所示。

自适应纠错标签损失函数用来标注训练过程中的错误负样本,降低抽样偏差导致的精度下降,如式(10)所示。

用于模型训练的总损失函数为

$$L = L_{eq} + \lambda L_a. \quad (11)$$

式中:  $\lambda$  为用来平衡  $L_{eq}$  和  $L_a$  的超参数。

## 3 实验部分

### 3.1 数据集

本文在 UCMerced 和 EuroSAT 数据集上进行实验,数据集包含多种不同的遥感图像,具有不同的场景和地物类型。UCMerced 包含 21 种不同的地物类型,每个类别共有 100 张图像,每张图像的大小为  $256 \times 256$  像素,如图 2 所示。EuroSAT 包含 10 种不同的地物类型,每个类别有 2 000~3 000 张图像,每张图像大小为  $64 \times 64$  像素,如图 3 所示。

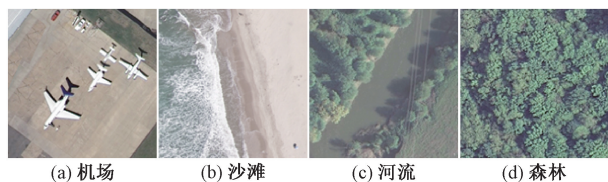


图2 UCMerced 示例图像

Figure 2 Example images of UCMerced

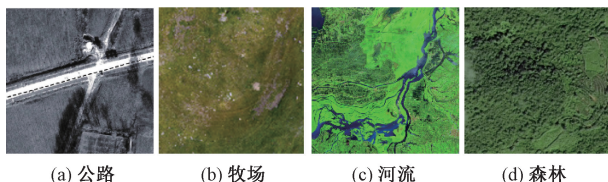


图3 EuroSAT 示例图像

Figure 3 Example images of EuroSAT

### 3.2 实验设置与评价指标

本文使用标准的 ResNet50 作为骨干网络,实验所用 GPU 为 TeslaP40,PyTorch 版本号为 1.8.1。

对于网络训练,使用 Adam 优化器,学习率为 0.000 3,使用余弦衰减策略调整学习速率,在 UC-Merced 上设置  $batch\ size$  为 64,在 EuroSAT 上设置

$batch\ size$  为 256。子量化器的数量  $m = \{4, 8, 12, 16\}$ ,将子量化器中码字的数量固定为  $K = 2^4$ 。遥感图像特征经过乘积量化后的量化维度  $D = m \log_2 K$ 。通过调整码本数量  $m$  来获得  $\{16, 32, 48, 64\}$  位二进制代码。模型中的超参数  $\phi_g = 5.0$ ,  $\lambda = 1$ ,  $\phi_a = 0.5$ ,  $\phi_{eq} = 0.5$ 。

本文使用  $mAP$  评估检索性能。 $mAP$  表示对于给定的查询  $q$ ,计算其平均精度  $AP$ ,最后计算所有查询图像的  $AP$  的平均值得到  $mAP$ 。平均精度  $AP$  为

$$AP = \frac{1}{R} \sum_{r=1}^{N'} P(r) y_r. \quad (12)$$

式中:  $N'$  为返回的图像数量;  $R$  表示相关图像的数量;  $P(r)$  表示返回图像的精确度;  $y_r$  表示第  $r$  个返回的图像是否与图像相关。

### 3.3 对比实验

在 EuroSAT 和 UCMerced 数据集上将 CQSC 与多种先进的方法进行比较。其中,有 3 种专门为遥感图像检索设计的无监督哈希方法,包括基于内核的无监督哈希 (KULSH)<sup>[16]</sup>、概率潜在语义哈希 (PLSH)<sup>[17]</sup> 和部分随机哈希 (PRH)<sup>[18]</sup>;还有 3 种在计算机视觉中广泛使用的自然图像哈希方法,即局部敏感哈希 (LSH)<sup>[19]</sup>、迭代量化 (ITQ)<sup>[20]</sup> 和双瓶颈哈希 (TBH)<sup>[21]</sup>。

表 1 展示了在 UCMerced 数据集上不同长度哈希码下的  $mAP@20$ 。在 16 bits 的哈希码长度下,  $mAP@20$  提升 3.73 百分点;在 32 bits 的哈希码长度下,相对于 PLSH,提升了 4.93 百分点;当哈希码长度达到 48 bits 时,相对于低位数哈希码情况下的结果,依旧提升 1.97 百分点,仍然保持较高的性能;当哈希码长度达到 64 bits 时,  $mAP@20$  略低于 PLSH,说明随着编码长度的增加,数据在高维空间中的分布变得更加稀疏,使得学习有效特征变得更加困难。同时,编码长度的增加会引入更多的冗余信息,也会影响遥感图像检索性能的提升。在 UC-Merced 数据集上的实验结果表明,本文方法的  $mAP@20$  比 PLSH 方法综合提高了 2.52 百分点。

表 2 为展示了在 EuroSAT 数据集上不同长度哈希码下的  $mAP@100$ 。在 16 bits 和 32 bits 的哈希码长度下, CQSC 略低于 PLSH 的  $mAP$ ;当哈希码长度达到 48 bits 时,相较于效果最好的 PLSH 精度提升 0.78 百分点;当哈希码长度达到 64 bits 时,PLSH 的精度有所下降,但是 CQSC 的精度依然提升,保持较高的性能水平, CQSC 的  $mAP@100$  比 PLSH 提升了 3.83 百分点。这说明 PLSH 在编码长度较短的时

候呈现了检索优势,但是随着编码的增长,PLSH 检索性能并没有持续提升。

表 1 在 UCMerced 上  $mAP@20$  的对比结果  
Table 1 Comparison results of  $mAP@20$  on UCMerced

方法	$mAP@20/\%$			
	16 bits	32 bits	48 bits	64 bits
LSH <sup>[19]</sup>	32.44	38.58	45.48	51.67
ITQ <sup>[20]</sup>	42.38	45.99	47.28	47.49
KULSH <sup>[16]</sup>	53.68	55.37	58.23	64.58
PRH <sup>[18]</sup>	55.39	59.45	61.89	67.77
PLSH <sup>[17]</sup>	62.28	65.35	70.44	73.07
CQSC	66.01	70.28	72.41	72.54

表 2 在 EuroSAT 上  $mAP@100$  的对比结果  
Table 2 Comparison results of  $mAP@100$  on EuroSAT

方法	$mAP@100/\%$			
	16 bits	32 bits	48 bits	64 bits
LSH <sup>[19]</sup>	32.24	40.59	46.42	50.90
ITQ <sup>[20]</sup>	60.21	65.07	66.72	67.68
KULSH <sup>[16]</sup>	70.10	74.10	—	—
TBH <sup>[21]</sup>	64.58	69.01	70.32	68.65
PLSH <sup>[17]</sup>	77.04	78.27	79.12	77.05
CQSC	72.54	77.01	79.90	80.88

3.4 消融实验

为证明 CQSC 的可行性,在 EuroSAT 和 UC-Merced 上进行消融实验,将哈希码的长度固定为 32 位,保持其他参数固定不变,分别进行  $epoch$  为 2 000 的实验,结果如表 3 所示。由表 3 可以看出,在 UC-Merced 数据集中,没有使用交叉量化学习策略和自适应纠错标签的情况下, $mAP@20$  为 67.67%,这表明模型在没有这两个模块的情况下,对于信息检索的效果相对较低。仅使用交叉量化学习策略的情况下, $mAP@20$  为 70.07%;而仅使用自适应纠错标签的情况下, $mAP@20$  为 68.68%,这表明交叉量化学习策略和自适应纠错标签的使用对于提升模型性能是有益的。最后,同时使用交叉量化学习和自适应纠错标签, $mAP@20$  为 70.28%,这表明两个模块的结合对于模型的性能有协同效应,能够进一步提升自监督遥感图像检索的准确性。

在 EuroSAT 数据集中,没有使用交叉量化学习策略和自适应纠错标签的情况下 CQSC 的  $mAP@20$  只有 70.41%,当分别采用交叉量化学习策略和自适应纠错标签时,提升了 2.19 个百分点和 4.77 百分点,这证明 CQSC 是可行的。当采用交叉量化学习策略和自适应纠错标签时,在 EuroSAT 数据集上表现最好, $mAP@20$  达到 77.01%。

表 3 消融实验  
Table 3 Ablation experiment

交叉 量化	自适应 纠错标签	$mAP@20/\%$	
		UCMerced 数据集	EuroSAT 数据集
		67.67	70.41
	✓	68.68	72.60
✓		70.07	75.18
✓	✓	70.28	77.01

注:✓表示采用该项策略。

综合分析,交叉量化学习策略和自适应纠错标签对模型的性能均有积极影响,同时使用的情况下,能够达到最佳的  $mAP$  分数。

3.5 参数敏感性分析

在 UCMerced 上的敏感性分析结果如表 4 所示。为证实验结果的最佳,首先,在 UCMerced 数据集中测试自适应纠错标签中控制量化分布的参数  $\phi_a$  对实验的影响,取值为 0 到 1,将哈希码的长度固定为 32 位,保持  $\phi_{eq}$  和其他超参数不变的情况下,训练 2 000 轮次。从表 4 中可以看出,当  $\phi_a$  为 0.1 时, $mAP@20$  为 64.43%;当  $\phi_a$  为 0.3 时, $mAP@20$  提高到 69.12%;当  $\phi_a$  为 0.5 时,进一步提高到 70.28%;当  $\phi_a$  为 0.7 时, $mAP@20$  降低到 67.8%。随着  $\phi_a$  的增加,平均准确度逐渐提高,说明适当增加自适应纠错标签的权重可以提升模型的性能。当  $\phi_a$  达到 0.5 时, $mAP@20$  达到 70.28%,这表明此时自适应纠错标签的权重达到一个良好的平衡,能够最大化地改善模型的性能。

表 4 在 UCMerced 上的敏感性分析实验结果  
Table 4 Experimental results of sensitivity analysis on UCMerced

参数取值	$mAP@20/\%$	参数取值	$mAP@20/\%$
$\phi_a=0.1$	64.43	$\phi_{eq}=0.1$	64.31
$\phi_a=0.3$	69.12	$\phi_{eq}=0.3$	69.09
$\phi_a=0.5$	70.28	$\phi_{eq}=0.5$	70.28
$\phi_a=0.7$	67.80	$\phi_{eq}=0.7$	64.39

其次,对交叉量化损失函数的非负温度参数  $\phi_{eq}$  进行最佳值测定,取值范围从 0 到 1,保持  $\phi_a$  和其他超参数不变。结果发现,当  $\phi_{eq}$  的值为 0.1 时, $mAP@20$  为 64.31%;当  $\phi_{eq}$  的值为 0.3 时, $mAP@20$  提高到 69.09%;当  $\phi_{eq}$  的值为 0.5 时, $mAP@20$  达到最高点,为 70.28%;然而,当  $\phi_{eq}$  的值增加到 0.7 时, $mAP@20$  却下降至 64.39%。

最后,将哈希码的长度固定为 32 bits,保持其他参数不变的情况下,训练 1 000 轮次,对平衡交叉量化损失和自适应纠错标签损失的参数  $\lambda$  进行

最佳值分析。表 5 为参数  $\lambda$  在 UCMerced 数据集上  $mAP@20$  和 EuroSAT 数据集上  $mAP@100$  的分析结果。

表 5  $\lambda$  的敏感性实验结果  
Table 5 Sensitivity test results of  $\lambda$

$\lambda$	$mAP@20/\%$	$mAP@100/\%$
0.1	66.57	71.52
0.5	67.58	71.38
1.0	70.28	77.01
1.5	68.44	76.01

由表 5 可以看出,随着自适应纠错损失的所占比重提升,模型能够获得更好的检索性能。当  $\lambda$  从 0.1 增加到 0.5 时,在 UCMerced 数据集上检索精度提升了 1.01 百分点,这表明自适应纠错损失在总损失中的权重较小,对模型的指导作用不足以产生大的变化。当  $\lambda$  继续增加到 1 时,相较于  $\lambda$  为 0.5 时  $mAP@20$  提升了 2.70 百分点,  $mAP@100$  提升了 5.63 百分点,这表明自适应纠错标签开始对模型产生了较强的影响,改善了检索性能。当  $\lambda$  增加到 1.5 时,性能有所下降,这表明自适应纠错标签占据过大的比重时会导致模型过度注意错误负样本而忽略了交叉量化损失,从而破坏了损失函数之间的平衡,影响模型的总体性能。

4 结论

本文提出了一种基于交叉量化和样本校正的自监督遥感图像检索模型(CQSC),根据特征空间中的样本间的相邻关系,对相似性进行加权作为自适应纠错标签减少训练样本中的错误负样本。通过交叉策略和乘积量化,高效压缩数据的同时,提升自监督遥感图像检索的准确率。

参考文献:

[1] 张艺琨,唐雁,陈强. 基于多特征融合的三维模型检索[J]. 郑州大学学报(工学版), 2019, 40(1): 1-6.  
ZHANG Y K, TANG Y, CHEN Q. 3D model retrieval method based on multiple feature fusion[J]. Journal of Zhengzhou University (Engineering Science), 2019, 40(1): 1-6.

[2] 冯孝鑫,王子健,吴奇. 基于三元采样图卷积网络的半监督遥感图像检索[J]. 电子与信息学报, 2023, 45(2): 644-653.  
FENG X X, WANG Z J, WU Q. Semi-supervised learning remote sensing image retrieval method based on triplet sampling graph convolutional network[J]. Journal of Electronics & Information Technology, 2023, 45(2): 644-

653.

[3] JING W P, XU Z K, LI L H, et al. Deep unsupervised weighted Hashing for remote sensing image retrieval[J]. Journal of Database Management, 2022, 33(2): 1-19.

[4] FENG C, PATRAS I. Adaptive soft contrastive learning [C]//The 26th International Conference on Pattern Recognition (ICPR). Piscataway: IEEE, 2022: 2721 - 2727.

[5] CHUANG C Y, ROBINSON J, YEN-CHEN L, et al. Debaised contrastive learning [C] //Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM, 2020: 8765-8775.

[6] AI L F, TAO Y, CHENG H J, et al. Accumulative quantization for approximate nearest neighbor search[J]. Computational Intelligence and Neuroscience, 2022, 2022: 4364252.

[7] SUN Y X, YE Y M, LI X T, et al. Unsupervised deep Hashing through learning soft pseudo label for remote sensing image retrieval [J]. Knowledge-Based Systems, 2022, 239: 107807.

[8] CHEN B, FENG Y, DAI T, et al. Adversarial examples generation for deep product quantization networks on image retrieval[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(2): 1388-1404.

[9] ZHOU H Y, QIN Q B, HOU J K, et al. Deep global semantic structure-preserving Hashing via corrective triplet loss for remote sensing image retrieval[J]. Expert Systems with Applications, 2024, 238: 122105.

[10] LIANG Y, ZHANG S L, LI K L, et al. Unleashing the full potential of product quantization for large-scale image retrieval[EB/OL]. (2023-10-09) [2024-08-22]. <https://paperswithcode.com/paper/unleashing-the-full-potential-of-product>.

[11] TANG J Y, WANG D L, TONG X C, et al. Learning from self-supervised features for Hashing-based remote sensing image retrieval [J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 2501705.

[12] HUYNH T, KORNBLITH S, WALTER M R, et al. Boosting contrastive self-supervised learning with false negative cancellation[C]//2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2022: 986-996.

[13] DWIBEDI D, AYTAR Y, TOMPSON J, et al. With a little help from my friends: nearest-neighbor contrastive learning of visual representations[EB/OL]. (2021-10-07) [2024-08-22]. <http://arxiv.org/abs/2104.14548>.

[14] 何悦,陈广胜,景维鹏,等. 基于深度多相似性哈希方法的遥感图像检索[J]. 计算机工程, 2023, 49(2): 206-212.

HE Y, CHEN G S, JING W P, et al. Remote sensing image retrieval based on deep multi-similarity Hashing method[J]. Computer Engineering, 2023, 49(2): 206–212.

[15] 吴刚, 葛芸, 储珺, 等. 面向遥感图像检索的级联池化自注意力研究[J]. 光电工程, 2022, 49(12): 220029.

WU G, GE Y, CHU J, et al. Cascade pooling self-attention research for remote sensing image retrieval[J]. Opto-Electronic Engineering, 2022, 49(12): 220029.

[16] DEMIR B, BRUZZONE L. Hashing-based scalable remote sensing image search and retrieval in large archives[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(2): 892–904.

[17] FERNANDEZ-BELTRAN R, DEMIR B, PLA F, et al. Unsupervised remote sensing image retrieval using probabilistic latent semantic Hashing[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 18(2): 256–260.

[18] LI P, REN P. Partial randomness hashing for large-scale remote sensing image retrieval[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(3): 464–468.

[19] SLANEY M, CASEY M. Locality-sensitive Hashing for finding nearest neighbors lecture notes[J]. IEEE Signal Processing Magazine, 2008, 25(2): 128–131.

[20] GONG Y C, LAZEBNIK S, GORDO A, et al. Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(12): 2916–2929.

[21] SHEN Y M, QIN J, CHEN J X, et al. Auto-encoding twin-bottleneck Hashing[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 2815–2824.

Self-supervised Remote Sensing Image Retrieval Based on  
Cross-quantization and Sample Correction

PAN Lili<sup>1</sup>, QU Dongliang<sup>2</sup>, YIN Jingjing<sup>2</sup>, MA Xueqiang<sup>1</sup>

(1. School of Computer Science and Mathematics, Central South University of Forestry and Technology, Changsha 410000, China; 2. School of Electronic Information and Physics, Central South University of Forestry and Technology, Changsha 410000, China)

**Abstract:** In view of the fact that the self-supervised remote sensing image retrieval model by using the sample pair for learning due to the lack of labels, resulted in sampling bias and affecting the accuracy of image representation, a self-supervised remote sensing image retrieval model based on cross-quantization and sample correction (CQSC) was proposed. Firstly, in order to reduce the load of data storage and processing, the mapping layer and product quantization in traditional contrast learning were combined to compress the high-dimensional image data and improve the retrieval efficiency. Secondly, the cross-learning strategy was used to maximize the cross-similarity before and after the feature mapping in the retrieval model, and the feature generation ability and retrieval accuracy of the model were enhanced. Finally, design adaptive correction labels to annotate training samples, correct erroneous negative samples during training and address sampling bias caused by missing labels in self-supervised remote sensing image retrieval. Experiments on UCMerced and EuroSAT datasets showed that compared with PLSH, *mAP@20* of CQSC improved by 2.52 percentage points on average on UCMerced, and *mAP@100* of CQSC improved by 3.83 percentage points on EuroSAT with 64bits.

**Keywords:** remote sensing image retrieval; contrastive learning; product quantization; cross-learning