

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

基于深度度量学习的卫星云图检索

金柱璋, 方旭源, 黄彦慧, 尹曹谦, 金炜

引用本文:

金柱璋, 方旭源, 黄彦慧, 等. 基于深度度量学习的卫星云图检索[J]. 光电工程, 2022, 49(4): 210307.

Jin Z Z, Fang X Y, Huang Y H, et al. Satellite cloud image retrieval based on deep metric learning[J]. *Opto-Electron Eng*, 2022, 49(4): 210307.

<https://doi.org/10.12086/oe.2022.210307>

收稿日期: 2021-09-23; 修改日期: 2022-01-18

相关论文

结合遥感卫星及深度神经决策树的夜间海雾识别

李涛, 金炜, 符冉迪, 李纲, 尹曹谦

光电工程 2022, 49(9): 220007 doi: [10.12086/oe.2022.220007](https://doi.org/10.12086/oe.2022.220007)

基于深度学习的门机抓斗检测方法

张文明, 刘向阳, 李海滨, 李雅倩

光电工程 2021, 48(1): 200062 doi: [10.12086/oe.2021.200062](https://doi.org/10.12086/oe.2021.200062)

基于深度学习检测器的多角度人脸关键点检测

赵兴文, 杭丽君, 宫恩来, 叶锋, 丁明旭

光电工程 2020, 47(1): 190299 doi: [10.12086/oe.2020.190299](https://doi.org/10.12086/oe.2020.190299)

面向遥感图像检索的级联池化自注意力研究

吴刚, 葛芸, 储珺, 叶发茂

光电工程 2022, 49(12): 220029 doi: [10.12086/oe.2022.220029](https://doi.org/10.12086/oe.2022.220029)

更多相关论文见光电期刊集群网站 



光电工程
Opto-Electronic Engineering

<http://cn.oeejournal.org/oe>



OE_Journal



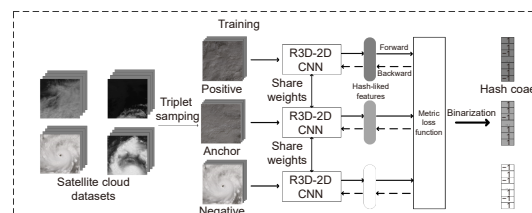
Website

DOI: 10.12086/oe.2022.210307

基于深度度量学习的卫星云图检索

金柱璋, 方旭源, 黄彦慧, 尹曹谦, 金 炜*

宁波大学信息科学与工程学院, 浙江 宁波 315211



摘要: 针对传统云图检索方法难于获得理想的检索精度且检索效率低的问题, 提出了一种基于深度度量学习的云图检索方法。首先设计了残差 3D-2D 卷积神经网络, 以提取云图的空间及光谱特征。鉴于传统基于分类的深度网络所提取的特征可能存在类内差异大、类间差异小的问题, 采用三元组训练网络, 依据云图之间的相似性将云图映射到度量空间中, 以使同类云图在嵌入空间中的距离小于非同类云图。在模型训练时, 通过对无损三元组损失函数增加正样本对间距离的约束, 改善了传统三元组损失的收敛性能, 提高了云图检索的精度。在此基础上, 通过哈希学习, 将度量空间中的云图特征变换成哈希码, 在保证检索精度的条件下提高了检索效率。实验结果表明, 在东南沿海云图数据集和北半球区域云图数据集上, 本文算法的平均精度均值 (mean average precision, mAP) 分别达到 75.14% 和 80.14%, 优于其他对比方法。

关键词: 深度学习; 度量学习; 三元组损失; 卫星云图检索

中图分类号: TP751

文献标志码: A

金柱璋, 方旭源, 黄彦慧, 等. 基于深度度量学习的卫星云图检索 [J]. 光电工程, 2022, 49(4): 210307

Jin Z Z, Fang X Y, Huang Y H, et al. Satellite cloud image retrieval based on deep metric learning[J]. *Opto-Electron Eng*, 2022, 49(4): 210307

Satellite cloud image retrieval based on deep metric learning

Jin Zhuzhang, Fang Xuyuan, Huang Yanhui, Yin Caoqian, Jin Wei*

Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, Zhejiang 315211, China

Abstract: Due to the traditional cloud image retrieval methods are difficult to obtain ideal retrieval accuracy and retrieval efficiency, a cloud image retrieval method based on deep metric learning is proposed. Firstly, a residual 3D-2D convolutional neural network is designed to extract spatial and spectral features of cloud images. Since the features extracted by the traditional classify-based deep network may have greater differences intra-classes than inter-classes, the triplet strategy is used to train the network, and the cloud images are mapped into the metric space according to the similarity between cloud images, so that the distance of similar cloud images in the embedded space is smaller than that of non-similar cloud images. In model training, the convergence performance of traditional triplet loss is improved and the precision of cloud image retrieval is increased by adding a constraint on the distance between positive sample pairs to the lossless triplet loss function. Finally, through hash learning, the cloud features in the metric space are transformed into hash codes, so as to ensure the retrieval accuracy and improve the retrieval efficiency. Experimental results show that the mean average precision (mAP) of the proposed

收稿日期: 2021-09-23; 收到修改稿日期: 2022-01-18

基金项目: 国家自然科学基金资助项目 (42071323); 宁波市公益类科技计划项目 (202002N3104)

*通信作者: 金炜, xyjw1969@126.com。

版权所有©2022 中国科学院光电技术研究所

algorithm is 75.14% and 80.14% for the southeast coastal cloud image dataset and the northern hemisphere cloud image dataset respectively, which is superior to other comparison methods.

Keywords: deep learning; metric learning; triplet loss; satellite cloud image retrieval

1 引言

气象卫星能从空中监测不同尺度的天气现象, 其所获取的卫星云图对于天气分析与预报均有重要的作用。近年来, 随着气象卫星技术的发展, 卫星云图的空间光谱分辨率和采集频率不断提高, 如何管理海量的卫星云图, 设计高效的云图检索系统成为困扰气象工作者的难题。卫星云图检索属于图像检索范畴, 主要包括基于文本的图像检索 (text-based image retrieval, TBIR) 和基于内容的图像检索 (content-based image retrieval, CBIR)。由于 TBIR 需耗费大量时间进行标注, CBIR 已成为图像检索的主流技术, 并在云图检索中得到应用。徐坚^[1]设计了一个基于颜色和纹理特征的云图检索系统, 提高了云图特征的表达能力; 上官伟^[2]通过提取云图的形状特征, 并对特征加权提高了检索精度; 李秀馨^[3]利用模糊 C 均值聚类分割出各云系后, 提取云系特征, 进一步改善了对于云系的表征能力; 徐晔晔^[4]结合 Otsu 和改进分形维数算法首先对云系进行分割, 并利用多特征融合实现了红外云图检索。虽然上述方法在云图检索性能上各有优势, 但均采用人工定义的底层特征来表征云图, 难以捕获云图蕴含的气象语义信息, 且特征维度往往较高, 算法的计算复杂度不容忽视。

随着人脑视觉感知机制研究的兴起, 深度学习在各领域异军突起, 特别是深度卷积神经网络 (deep convolutional neural networks, DCNN) 在图像分类^[5]、检索^[6]和分割^[7]等领域都得到了广泛的应用。区别于传统自然图像, 卫星云图具有多波段特点, 当使用传统二维卷积神经网络 (2D convolutional neural networks, 2D CNN) 提取特征时, 难以有效刻画云图的波段信息。受 Roy 等人在高光谱图像分类工作^[8-9]的启发, 本文通过构建残差三维-二维卷积神经网络 (residual 3D-2D convolutional neural network, R3D-2D CNN), 提出了一种基于深度度量学习的卫星云图检索方法。该方法首先利用三维卷积和二维卷积分别提取云图的光谱信息和空间信息, 以空谱联合特征捕捉的策略表征云图蕴含的气象语义信息。在此基础上, 针对部分云图在特征空间存在类内差异大类间差异小的问题, 采用深

度度量学习^[10-12] (deep metric learning, DML), 通过三元组训练方法^[10], 将云图映射到度量空间中, 使得样本特征间的距离满足检索的要求。同时, 鉴于传统三元组损失易使模型落入局部最优的情况, 本文对三元组损失函数进行改进, 在解决早熟收敛问题的同时, 进一步拉近相似云图在嵌入空间中的距离, 并引入哈希学习以汉明距离计算云图间的相似度, 从而在保证检索精度的情况下提高了检索效率。

2 数据获取及模型构建

2.1 数据集构建

Himawari-8 气象卫星是日本发射的新一代静止卫星, 其上搭载的成像仪 (Advanced Himawari Imager, AHI) 较上一代卫星在频带数量、空间分辨率和时间频率方面都有很大改善^[13]。该卫星具有 3 个可见光、3 个近红外和 10 个红外观测波段, 3 个可见光波段中心波长分别为 0.46 μm 、0.51 μm 和 0.64 μm , 3 个近红外波段中心波长分别为 0.96 μm 、1.6 μm 和 2.3 μm , 10 个红外波段中心波长分布于 3.9 μm ~13.3 μm 之间。其中, 可见光和近红外波段的云图数据以反射率形式呈现, 而红外波段以亮度温度形式呈现, 二者可以分别反映云层的厚度和云层的高度^[14]。如果将不同模态的卫星云图简单融合设计云图检索系统, 不仅难于提高检索的准确度, 而且将带来额外的计算复杂度。因此传统云图检索算法往往基于 3 个可见光波段进行设计, 这就忽略了能够提供如水/冰相态、粒子大小和光学厚度等云物理参数的其它波段^[13]的信息。为了综合利用不同波段的信息, 本研究选取 3 个可见光和 3 个近红外波段的云图数据开展检索系统的研究。

本文所选取的 Himawari-8 气象卫星云图由日本宇宙航空研究开发署的 P-Tree 系统所发布, 研究区域为东经 121°~125°, 北纬 28°~32°的中国东南沿海, 时间范围包括 2017 与 2018 年每日北京时间 10 点至 11 点。我们将全圆盘云图数据依据研究区域的经纬度进行截取, 生成大小为 81 pixel×81 pixel 的不同时刻的 6 波段云图后组成数据集, 并对各云图依据对应气象站的历史天气状况记录进行标注。数据集包括

1244 个时次的阴天云图、372 个时次的晴天云图、1219 个时次的雨天云图和 792 个时次的台风云图, 图 1 为研究区域不同天气状况的可见光波段 1 的云图示例。

在此基础上, 为了刻画完整的天气系统, 本文参考云图数据集 LSCIDMR^[15] 的构建方法, 以 1000 pixels×1000 pixels 大小的滑动窗口对 Himawari-8 的北半球云图进行截取, 并参照 LSCIDMR 的标注将所截取的云图标注为热带飓风、锋面、西风急流、雪等不同的天气系统; 为了控制算法的计算复杂度, 本文将 6 通道的 1000 pixels×1000 pixels 的云图缩小为 128 pixels×128 pixels; 为了使不同类别天气系统的云图数量基本均衡, 我们筛选了 632 个时次的锋面云图、628 个时次的西风急流云图、717 个时次的雪天云图和 698 个时次的热带飓风云图组成北半球区域云图数据集, 图 2 给出了不同天气系统可见光波段 1 的云图示例。下文的研究将分别基于东南沿海区域云图数据集和北半球区域云图数据集展开, 为了叙述简洁, 将它们分别简称为沿海云图数据集和北半球云图数据集。

2.2 模型构建

设卫星云图数据集为 $I = \{x_1, x_2, \dots, x_i, x_n\}$, n 代表数

据集样本的数量, x_i 为每时次的云图, 其对应的天气标签为 y_i 。为了实现云图检索系统, 本模型应用哈希函数将卫星云图编码为二值哈希码: $h: I \rightarrow \{-1, 1\}^{K \times n}$, 其中 K 为哈希码的位数。我们首先构建残差三维-二维卷积神经网络模型以刻画云图的空谱联合特征, 然后通过度量学习监督模型训练, 提取云图的哈希码特征。图 3 为算法流程图。

本模型采用两个时次的同类云图和一个时次的非同类云图组成三元组, 开展度量学习训练。三元组通过 R3D-2D CNN 生成近似哈希码, 经过二值化操作转换成哈希码后存储到哈希码库中。在测试阶段, 待检索云图的哈希码同云图库中历史云图的哈希码进行比对, 并通过计算汉明距离返回检索结果。

2.2.1 残差三维-二维卷积神经网络

为了提取卫星云图的空间及多波段光谱特征, 本文在二维卷积网络的基础上, 构建了如图 4 所示的残差 3D-2D 卷积神经网络。

从图中可知, 残差 3D-2D 卷积神经网络分为两个模块: 光谱特征学习模块 (spectral feature learning) 和空间特征学习模块 (spatial feature learning)。其中在光谱特征学习模块 (图 4 上半部分) 中, 为了加快模型

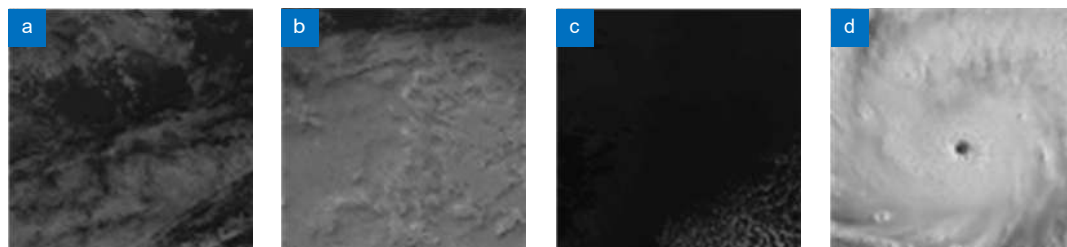


图 1 不同天气可见光波段 1 的云图。

(a) 阴天; (b) 雨天; (c) 晴天; (d) 台风

Fig. 1 Visible band 1 cloud images of different weather

(a) Cloudy; (b) Rainy; (c) Fair; (d) Typhoon

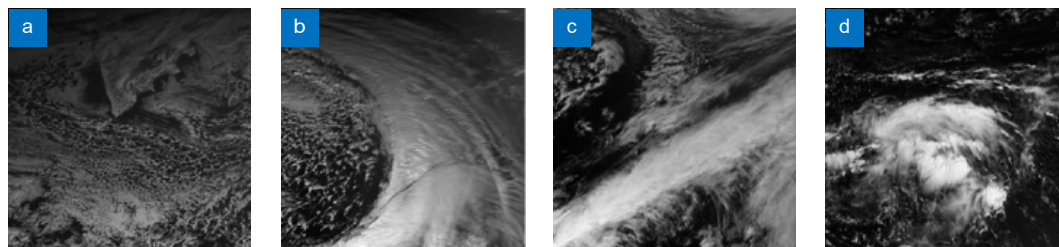


图 2 不同天气系统可见光波段 1 的云图。

(a) 雪; (b) 锋面; (c) 西风急流; (d) 热带飓风

Fig. 2 Visible band 1 cloud images of different weather systems.

(a) Snow; (b) Frontal surface; (c) Westerly jet; (d) Tropical cyclone

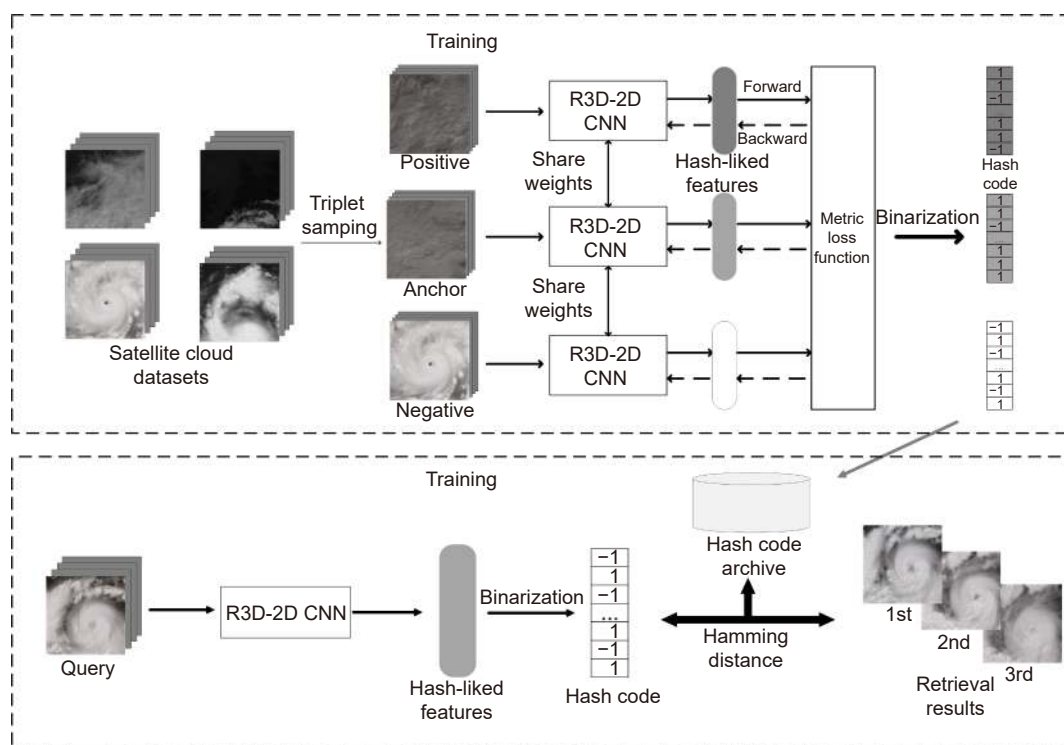


图3 算法流程图

Fig. 3 Overall algorithm flow chart

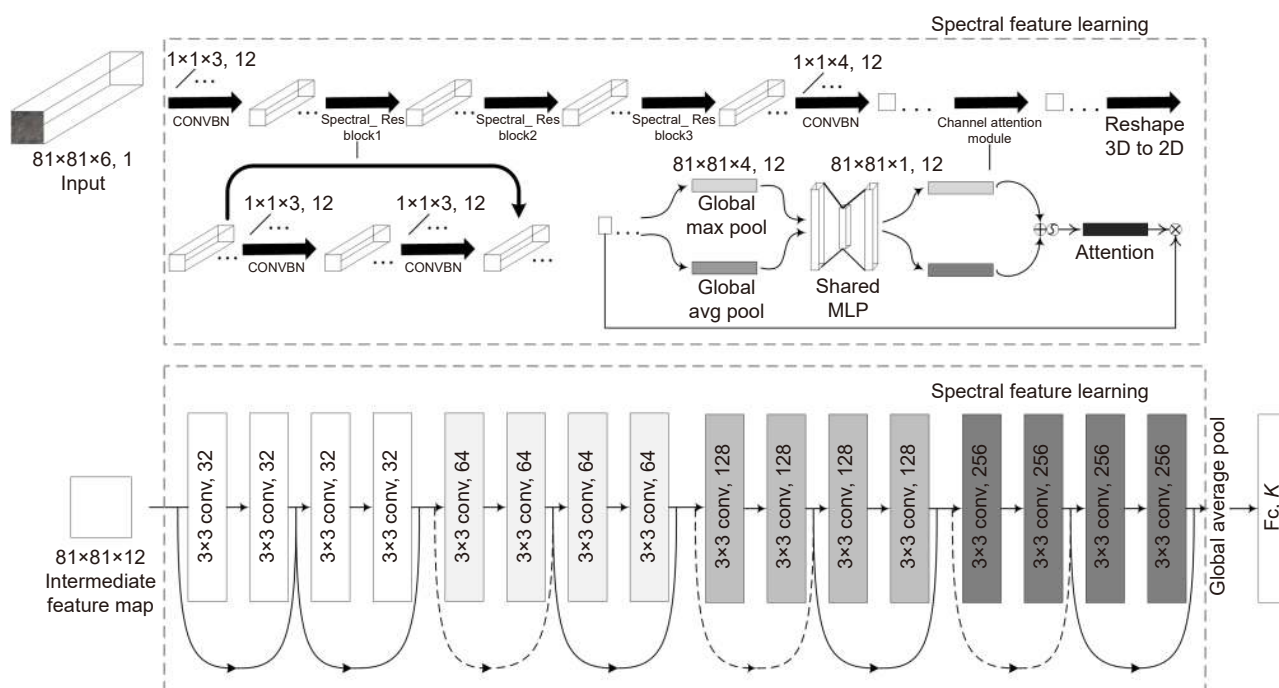


图4 残差3D-2D卷积神经网络

Fig. 4 Residual 3D-2D convolution neural network

收敛并克服网络退化, 引入了残差块 (Spectral_ResBlock)^[16] 的结构, 通过多个残差块, 不仅可以提取云图的深层光谱特征, 而且通过恒等映射, 模型训练时梯度能在高低层间迅速传递, 从而促进和规范化模型的训练过程。为有利于云图光谱特征的提取, 本文在模型中引入了通道注意力机制。首先通过全局平均池化 (global average pooling) 和全局最大池化 (global max pooling) 对特征图进行聚合, 之后将聚合的特征图输入一个共享的多层感知机 (multilayer perceptron, MLP) 来生成对应的特征向量, 特征向量按各元素相加并经过 sigmoid 激活后获得通道注意力权重向量。最后运用权重向量对原始特征图进行加权, 以细化特征图中各个通道的权重, 使模型关注有意义的通道而削弱对无意义通道的关注度。在提取云图光谱信息的基础上, 为了进一步刻画云图的空间信息, 该模型通过矩阵变维操作, 改变特征图的维度, 输入空间特征学习模块。在空间特征学习模块中 (图 4 下半部分), 本文以残差块的形式构建二维卷积来提取云图的空间信息。考虑到网络最终以哈希码的形式表示卫星云图, 我们采用全局平均池化整合特征图的全局空间信息, 并输入神经元个数为 K 的全连接层, 以双曲正切函数 (tanh) 作为全连接层的激活函数, 获得 K 位 $(-1,1)$ 的近似哈希码。

2.2.2 深度度量学习

由于采用传统交叉熵分类损失训练的深度学习难于全面刻画云图间的差异, 致使所提取的云图特征会存在类间差异小、类内差异大的情况^[17]。为了解决这一问题, 本文引入度量学习, 采用三元组训练将云图特征映射到度量空间, 使得在度量空间中同类云图间

的距离小于非同类云图。三元组 $t = \{(g^a, g^p, g^n)\}$ 由锚样本、正样本和负样本组成, 其中锚样本 g^a 和正样本 g^p 类别相同, 而负样本 g^n 则属于不同类别。在训练时, g^a 和 g^p 组成一组正样本对, 而 g^a 和 g^n 组成一组负样本对, 通过网络训练, 使得正样本对间距离缩小, 负样本对间距离扩大^[18], 如图 5 所示。

2.3 目标函数构建

传统三元组度量学习采用哈希码以汉明距离来计算样本间的距离, 造成使用三元组损失进行模型训练时梯度难于计算, 文献 [19] 对二进制约束进行连续松弛 (continuous relaxation), 将离散的哈希码转换成近似哈希码 (即连续的云图特征向量), 经过连续松弛后三元组损失如下:

$$L = \sum_{i=1}^D \max(0, \|f(g_i^a) - f(g_i^p)\|_2^2 - \|f(g_i^a) - f(g_i^n)\|_2^2 + \sigma), \quad (1)$$

式中: D 代表三元组的数目, g_i^a 、 g_i^p 、 g_i^n 为锚样本、正样本和负样本, 它们经过网络后所生成的近似哈希码分别为 $f(g_i^a)$ 、 $f(g_i^p)$ 、 $f(g_i^n)$, σ 为最小边际阈值。三元组损失能够抑制类内差异, 并扩大非同类云图间的距离。在模型训练时, 三元组损失相对于正样本、负样本和锚样本似哈希码的梯度分别为^[20]

$$\frac{\partial L}{\partial f(g_i^p)} = \sum_{i=1}^D 2 \cdot (f(g_i^p) - f(g_i^a)) \cdot \zeta(L > 0), \quad (2)$$

$$\frac{\partial L}{\partial f(g_i^n)} = \sum_{i=1}^D 2 \cdot (f(g_i^a) - f(g_i^n)) \cdot \zeta(L > 0), \quad (3)$$

$$\frac{\partial L}{\partial f(g_i^a)} = -\frac{\partial L}{\partial f(g_i^p)} - \frac{\partial L}{\partial f(g_i^n)}, \quad (4)$$

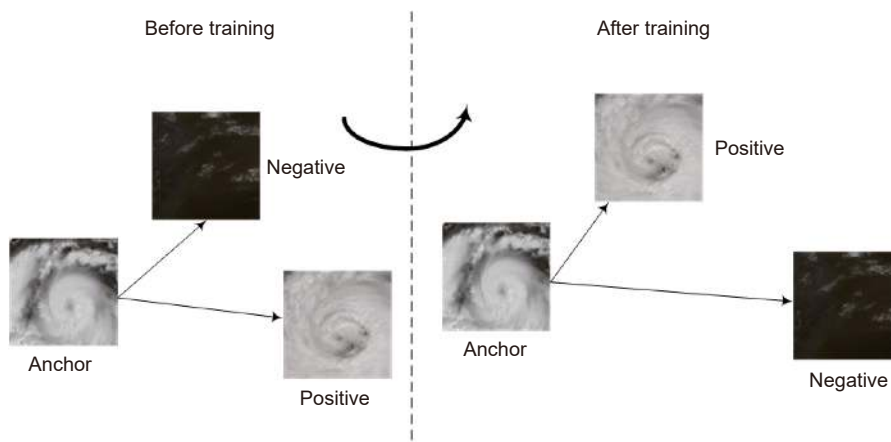


图 5 训练后正样本对间距离缩小, 负样本对间距离扩大

Fig. 5 After training, the distance of the anchor-positive decreases and the distance of the anchor-negative increases

其中 $\xi()$ 代表指示函数, 当三元组损失 L 大于 0 时为 1, 其他情况为 0。可以看出, 三元组损失所训练网络的性能依赖于阈值的设置, 当正样本对间的距离与阈值之和小于负样本对间的距离时, 梯度为 0, 模型参数无法得到更新, 模型陷入早熟收敛。为解决此问题, Arsenault 提出了无损三元组损失^[21]:

$$L_{\text{metric}} = \sum_{i=1}^D \left(-\ln \left(-\frac{\|f(g_i^a) - f(g_i^p)\|_2^2}{N} + 1 + \varepsilon \right) - \ln \left(-\frac{N - \|f(g_i^a) - f(g_i^p)\|_2^2}{N} + 1 + \varepsilon \right) \right), \quad (5)$$

式中: N 为哈希码位数的 4 倍即 $4K$, ε 是一个小正数以避免 $\ln 0$ 的出现。由于正样本对间的距离 $\|f(g_i^a) - f(g_i^p)\|_2^2 \in (0, 4K)$, 因此 $\frac{\|f(g_i^a) - f(g_i^p)\|_2^2}{N} \in (0, 1)$ 。利用此损失函数, 只有当锚样本与正样本间的距离接近 0 时, 损失函数才会达到最小值, 因此能解决传统三元组损失的早熟收敛问题。然而, 式 (5) 计算的三元组损失缺少对正样本对间距离的限制, 从而可能造成同类云图在嵌入空间中不够紧密^[22], 并降低模型的检索性能。为此, 我们在无损三元组损失的基础上对正样本对间的距离进行约束, 得到如下所示的损失函数:

$$L_{\text{metric}} = \sum_{i=1}^D \left(-\ln \left(-\frac{\|f(g_i^a) - f(g_i^p)\|_2^2}{N} + 1 + \varepsilon \right) - \ln \left(-\frac{N - \|f(g_i^a) - f(g_i^p)\|_2^2}{N} + 1 + \varepsilon \right) \right) + \lambda \sum_{i=1}^D \max(0, \|f(g_i^a) - f(g_i^p)\|_2^2 - \beta), \quad (6)$$

式中: λ 为平衡参数, 而 β 规定类内相似度的阈值。我们将卫星云图所对应的近似哈希码使用符号函数 (sign) 进行二值化处理, 生成云图对应的哈希码。在云图检索时, 通过待检索云图与云图库中历史云图的哈希码的比较, 即可返回检索结果。

3 实验结果与分析

本文实验的硬件环境为 Intel Core i5-10600KF CPU@4.10 GHz, 32 GB 内存, NVIDIA GeForce RTX 3060, 基于 Keras 和 tensorflow 框架构建深度网络模型。我们将沿海云图数据集和北半球云图数据集分别按 8:2 和 7:3 划分训练集和测试集。对每时次的云图旋转 90°、180°和 270°以扩增训练集, 尽量减轻网络

的过拟合问题。参照文献 [23] 的策略将锚样本与容易正样本&困难负样本 (easy positive & hard negative) 组成三元组, 采用自适应矩估计 (adaptive moment estimation, Adam) 优化器进行网络训练, 初始学习率设置为 10^{-5} , 目标函数中的超参数 λ 、 β 分别设置为 0.01 和 0.002。采用准确率 (precision, P) 和平均精度均值 (mean average precision, mAP) 两个指标对检索方法的性能进行评估^[24]:

$$P_{\text{Precision@L}} = \frac{1}{Q} \sum_{i=1}^Q \frac{m(x_i)}{L}, \quad (7)$$

$$mAP = \frac{\sum_{q=1}^Q AP(x_i)}{Q}, \quad (8)$$

式中: L 为检索时给定的返回云图数量, $m(x_i)$ 代表返回结果中与待检索云图 x_i 同类别的云图数量, Q 为待检索图像的总数, 为了表示简洁, 将 $P_{\text{Precision@L}}$ 统一简称为 $P@L$ 。式 (8) 中, AP 为每张云图的检索精度均值 (average precision, AP), 计算方法如下:

$$AP(x_i) = \frac{1}{M(x_i)} \sum_{j=1}^{M(x_i)} \frac{j}{L_j}, \quad (9)$$

式中: $M(x_i)$ 表示在云图数据集中与 x_i 同类别的云图数量, 在第 j 次检索时, 当给定返回的云图数量为 L_j 时, 系统刚好检索到 j 个时次的同类别云图。以下将通过实验分析损失函数和哈希码长度对模型性能的影响, 并将本文模型与其他检索算法进行比较。

3.1 损失函数对模型性能的影响

为了评价不同损失函数对于本模型的影响, 分别以三元组损失 (triplet loss, TL)、无损三元组损失 (lossless triplet loss, LTL) 和限制正样本对距离的无损三元组损失 (constrained lossless triplet loss, C-LTL) 作为目标函数对模型进行训练, 并比较了采用不同损失函数模型的检索性能。实验中, 哈希码长度设为 64 位, 三元组损失中的阈值设定为 0.2, 实验结果如表 1 和表 2 所示。

从表 1 和表 2 可以看出, 基于传统三元组损失函数的模型在两个数据集的检索实验中, 检索准确率与平均精度均值均不理想, 这可能是由于采用传统三元组损失函数时模型早熟收敛落入伪优造成的; 采用无损三元组损失的模型在检索准确率及平均精度均值指标上均有所改善, 在返回 20 张云图时, 在两个数据集上检索准确率分别提高了 3.49% 和 3.08%。而采用

表 1 沿海云图数据集中不同损失函数的模型检索性能

Table 1 The model retrieval performance of different loss functions in coastal cloud image dataset

Loss	P@5/%	P@20/%	mAP/%
TL	83.61	72.53	64.21
LTL	85.95	76.02	71.64
C-LTL	90.96	78.14	75.14

表 2 北半球云图数据集中不同损失函数的模型检索性能

Table 2 The model retrieval performance of different loss functions in the North hemisphere cloud image dataset

Loss	P@5/%	P@20/%	mAP/%
TL	80.57	79.47	71.23
LTL	87.70	82.55	73.07
C-LTL	85.20	85.63	80.14

限制正样本对距离的无损三元组损失函数模型, 由于对正样本对间的距离施加约束, 使得嵌入空间中同类云图的类簇更加紧密, 模型的检索性能得到了进一步的提高, 在两个数据集上 mAP 分别达到了 75.14% 和 80.14%。

3.2 哈希码长度对检索性能的影响

不同的哈希码长度会对检索系统的性能产生影响, 以沿海云图数据集为例, 将哈希码长度 K , 分别设定为 32、64、128 和 256 位, 开展检索性能的对比实验, 结果如图 6 所示。

从上图可以看出, 当 $K=32$ 时, 在返回 30 张云图时, 准确率仅为 68.69%, 虽然随着 K 值的增大, 检索性能会得到改善, 然而当 K 达到 64 后, 模型性能基本趋于稳定。从计算效率上看, $K=256$ 时检索一张云图约需 1.71 ms, 而 $K=64$ 时检索一张云图仅需 0.96 ms, 综合考虑检索性能与计算效率, 本文将哈希码长度设定为 64 位。

3.3 不同检索算法性能比较

为了验证本文算法的有效性, 在沿海云图数据集和北半球云图数据集上开展了与其它检索算法的对比实验, 对比算法包括: 1) 基于核的监督哈希 (kernel-based supervised Hashing, KSH)^[25]; 2) 基于深度学习的二进制哈希 (DLBHS)^[26]; 3) 基于度量学习的深度哈希 (metric-learning-based deep Hashing network, MiLan)^[18]; 4) 深度监督哈希 (deep supervised Hashing, DSH)^[19]。其中 KSH 方法采用三个可见光通道叠加后所提取的 GIST 特征^[27], DLBHS 方法使用 VGG16 作为特征提取器, 所有方法的哈希码长度均为 64 位, 实验结果如表 3 所示。

从上表可以看出, KSH 方法的检索性能最不理想, mAP 仅有 60% 左右, 这主要是由于 KSH 所采用的人工定义的特征难于捕捉云图的深层信息; 基于深度学习的 DLBHS、DSH 与 Milan 方法在获得图像的深层特征后, 再生成哈希码, 可以有效表征云图的深层信息, 因此检索性能有所改善。相较于本文算法,

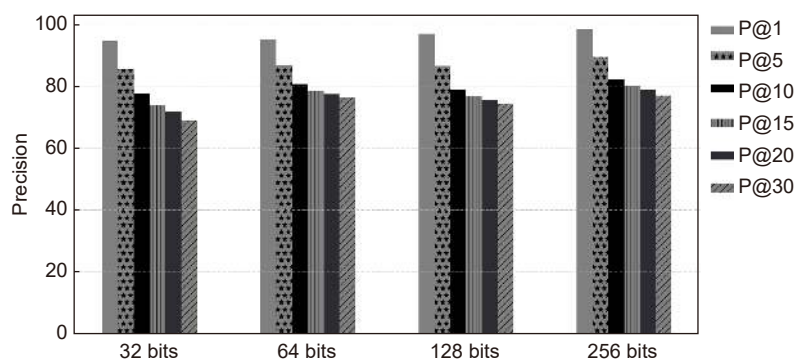


图 6 哈希码长度对模型的影响

Fig. 6 The effects of the hash code length on model performance

这些基于深度学习方法的检索性能不够理想,一方面是因为在特征提取阶段未能有效刻画云图的光谱信息;另一方面是因为在哈希码生成阶段所采用的相关度量损失函数不够优化。本文方法则通过 R3D-2D CNN 更好地表征了云图的空间和光谱信息,而且采用改进的无损三元组损失优化了嵌入空间中的样本分布,因此检索性能有了进一步的改善,在沿海云图数据集和北半球云图数据集上, mAP 分别达到 75.14% 和 80.14%。

为了直观展示本文方法的检索性能,给出了不同方法的检索实例。对于沿海云图数据集,选取 2018 年 1 月 30 日的多云天气云图 (cloudy) 作为待检索云

图,检索结果如图 7 所示。

从图 7 可以看出, KSH 方法的检索结果不够理想, Top7 中包含了雨天云图 (rainy); DLBHS、DSH 与 Milan 方法虽然返回的云图均属于多云天气,但没能有效返回相邻时次的相似云图,从视觉表现上看,与待检索云图相似度的排序也不合理,而且返回结果中部分云图的云团形状与纹理表现和待检索云图差异较大;本文方法不仅有效返回了相邻时次的相似云图,而且返回结果大都与待检索云图具有一致的形状与纹理表现。对于北半球云图数据集,选取了 2019 年 4 月 21 日的西风急流云图 (Westerly Jet) 进行检索实验,结果如图 8 所示。

表 3 各数据集在不同方法下的检索准确度

Table 3 Comparison of retrieval performance between different retrieval methods

Dataset	Methods	mAP/%	P@5/%	P@10/%	P@20/%	P@30/%
沿海云图数据集	KSH	60.10	73.62	71.73	67.76	66.30
	DLBHS	72.44	76.19	75.28	74.35	72.15
	MiLan	68.09	73.38	72.54	71.37	70.92
	DSH	68.61	85.76	81.17	76.89	73.24
	Proposed	75.14	90.96	82.41	78.14	76.68
北半球云图数据集	KSH	60.27	68.53	68.43	67.41	66.92
	DLBHS	78.13	84.65	83.70	83.00	82.71
	MiLan	74.97	81.33	80.64	80.16	79.85
	DSH	70.22	69.86	72.80	73.40	73.90
	Proposed	80.14	85.20	85.84	85.63	84.90

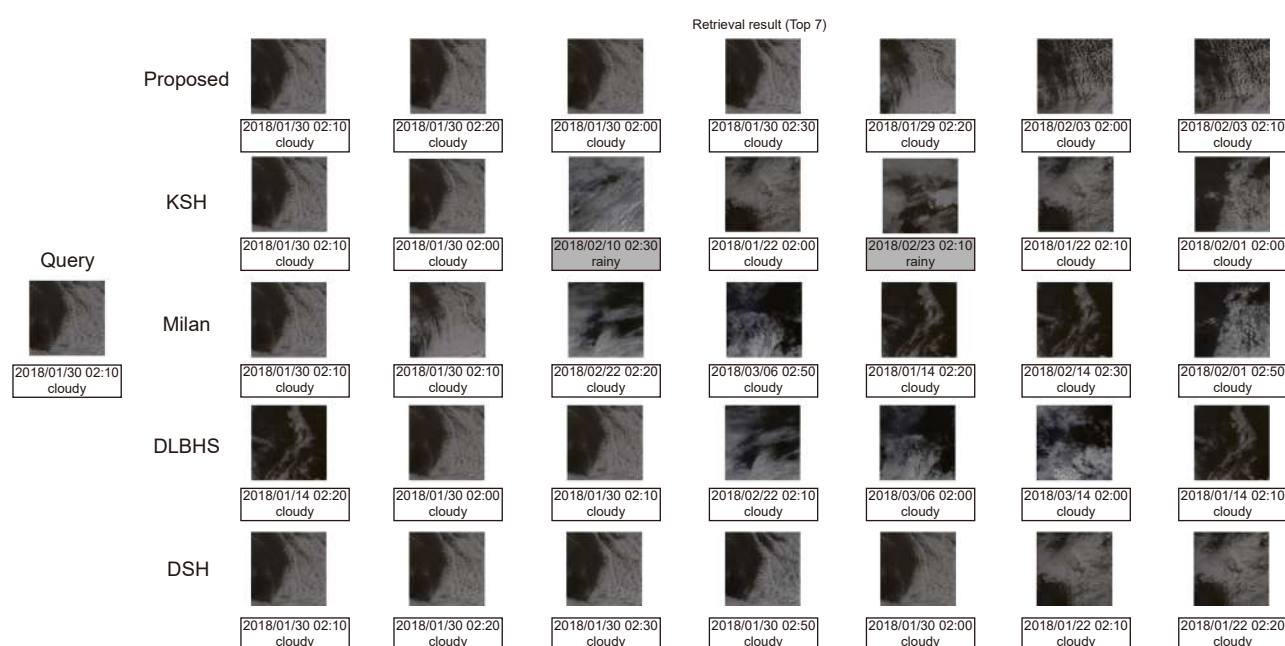


图 7 多云天气云图实例检索结果

Fig. 7 Retrieval results of cloudy weather image

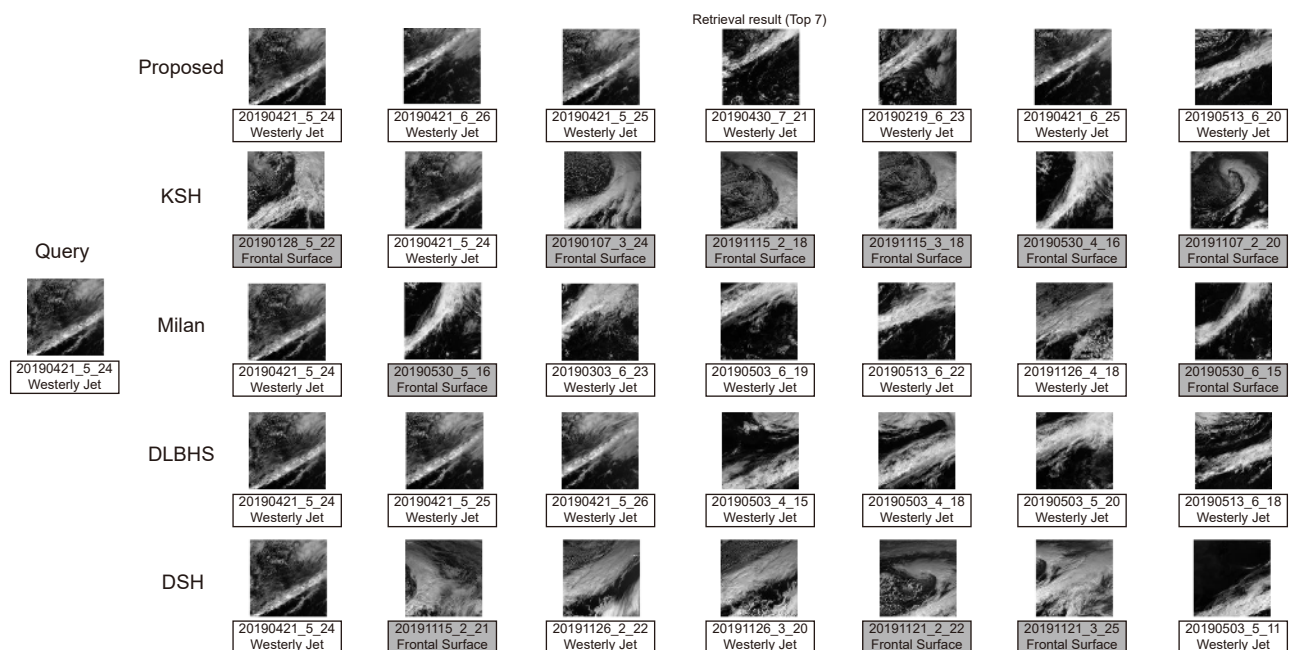


图 8 西风急流云图实例检索结果

Fig. 8 Retrieval results of westerly jet cloud image

从图 8 可以看出, KSH 方法的表现最不理想, 返回的云图大都为锋面云图 (frontal surface); DSH 与 Milan 方法的返回结果也均包含与待检索云图非同类的锋面云图, 虽然 DLBHS 方法的返回结果均为西风急流云图, 但从视觉相似度上看, 本文方法的返回结果与待检索云图的相似度排序更加合理。

4 结 论

为了提取多波段卫星云图的光谱和空间信息, 本文通过残差 3D-2D 卷积神经网络来表征卫星云图, 同时引入度量学习, 采用三元组训练得到云图的哈希码特征, 实现云图检索。在模型训练时, 通过对无损三元组损失函数增加正样本对间距离的约束, 改善了传统三元组损失的收敛性能。实验结果表明, 在东南沿海云图数据集和北半球区域云图数据集上, mAP 分别达到了 75.14% 和 80.14%, 优于其它对比方法; 下一步将研究如何更好地利用卫星云图的红外与可见光信息, 以提高云图检索方法的性能。

利益冲突: 所有作者声明无利益冲突

参考文献

- [1] Xu J. Content based cloudy satellite images retrieval system design and practice[D]. Nanjing: Hohai University, 2004.
徐坚. 基于内容的卫星云图检索系统设计与实现[D]. 南京: 河海大

- 学, 2004.
- [2] Shangguan W. Research of content-based satellite cloud image processing and information retrieval technology[D]. Harbin: Harbin Engineering University, 2008.
上官伟. 基于内容的卫星云图处理与信息检索技术研究[D]. 哈尔滨: 哈尔滨工程大学, 2008.
- [3] Li X X. Research on the technology of content-based satellite cloud image retrieval[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2013.
李秀馨. 基于内容的卫星云图检索技术研究[D]. 南京: 南京航空航天大学, 2013.
- [4] Xu Y Y. Infrared cloud image retrieval technology research based on multi-feature[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2014.
徐晔晔. 基于多特征的红外云图检索技术研究[D]. 南京: 南京航空航天大学, 2014.
- [5] Zhang C J, Li G D, Lei R M, et al. Deep feature aggregation network for hyperspectral remote sensing image classification[J]. *IEEE J Select Top Appl Earth Observ Remote Sens*, 2020, **13**: 5314–5325.
- [6] Fang J S, Fu H Z, Liu J. Deep triplet hashing network for case-based medical image retrieval[J]. *Med Image Anal*, 2021, **69**: 101981.
- [7] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//*Proceedings of the 2015 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015: 234–241.
- [8] Roy S K, Krishna G, Dubey S R, et al. HybridSN: exploring 3-D –2-D CNN feature hierarchy for hyperspectral image classification[J]. *IEEE Geosci Remote Sens Lett*, 2020, **17**(2): 277–281.
- [9] Feng F, Wang S T, Wang C Y, et al. Learning deep hierarchical spatial-spectral features for hyperspectral image classification based on residual 3D-2D CNN[J]. *Sensors*, 2019, **19**(23): 5276.
- [10] Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person re-identification[Z]. arXiv: 1703.07737, 2017. <https://doi.org/10.12086/oe.2022.210307>

- [org/10.48550/arXiv.1703.07737](https://doi.org/10.48550/arXiv.1703.07737).
- [11] Kaya M, Bilge H Ş. Deep metric learning: A survey[J]. *Symmetry*, 2019, 11(9): 1066.
 - [12] Schroff F, Kalenichenko D, Philbin J. FaceNet: a unified embedding for face recognition and clustering[C]//*Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 815–823.
 - [13] Bessho K, Date K, Hayashi M, et al. An introduction to Himawari-8/9 —Japan's new-generation geostationary meteorological satellites[J]. *J Meteor Soc Jpn Ser II*, 2016, 94(2): 151–183.
 - [14] Yu F, Chen W M. Research on the cloud classification for the bi-spectrum cloud picture[J]. *J Nanjing Inst Meteor*, 1994, 17(1): 117–124.
郁凡, 陈渭民. 双光谱云图的云分类探讨[J]. *南京气象学院学报*, 1994, 17(1): 117–124.
 - [15] Bai C, Zhang M J, Zhang J L, et al. LSCIDMR: large-scale satellite cloud image database for meteorological research[J]. *IEEE Trans Cybern*, 2021.
 - [16] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770–778.
 - [17] Huang K K, Ren C X, Liu H, et al. Hyperspectral image classification via discriminative convolutional neural network with an improved triplet loss[J]. *Pattern Recogn*, 2021, 112: 107744.
 - [18] Roy S, Sangineto E, Demir B, et al. Metric-learning-based deep hashing network for content-based retrieval of remote sensing images[J]. *IEEE Geosci Remote Sens Lett*, 2021, 18(2): 226–230.
 - [19] Liu H M, Wang R P, Shan S G, et al. Deep supervised hashing for fast image retrieval[C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 2064–2072.
 - [20] Kim S, Seo M, Laptev I, et al. Deep metric learning beyond binary supervision[C]//*Proceedings of the 2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 2283–2292.
 - [21] Arsenault Marc-Olivier. Lossless Triplet loss[EB/OL]. (2018-02-15). <http://coffeeanddata.ca/lossless-triplet-loss>.
 - [22] Cheng D, Gong Y H, Zhou S P, et al. Person Re-identification by multi-channel parts-based CNN with improved triplet loss function[C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1335–1344.
 - [23] Xuan H, Stylianou A, Pless R. Improved embeddings with easy positive triplet mining[C]//*Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision*, 2020: 2463–2471.
 - [24] Liao L F, Li Z M, Zhang S S. Image retrieval method based on deep residual network and iterative quantization hashing[J]. *J Comput Appl*, 2021. doi: 10.11772/j.issn.1001-9081.2021071135.
廖列法, 李志明, 张赛赛. 基于深度残差网络的迭代量化哈希图像检索方法[J]. *计算机应用*, 2021. doi: 10.11772/j.issn.1001-9081.2021071135.
 - [25] Liu W, Wang J, Ji R R, et al. Supervised hashing with kernels[C]//*Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012: 2074–2081.
 - [26] Lin K, Yang H F, Hsiao J H, et al. Deep learning of binary hash codes for fast image retrieval[C]//*Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015: 27–35.
 - [27] Torralba A, Murphy K P, Freeman W T, et al. Context-based vision system for place and object recognition[C]//*Proceedings of the Ninth IEEE International Conference on Computer Vision*, 2003: 273–280.

作者简介



金柱璋 (1997-), 男, 硕士研究生, 研究方向: 深度学习、图像检索、图像处理。

E-mail: fatejzz@foxmail.com

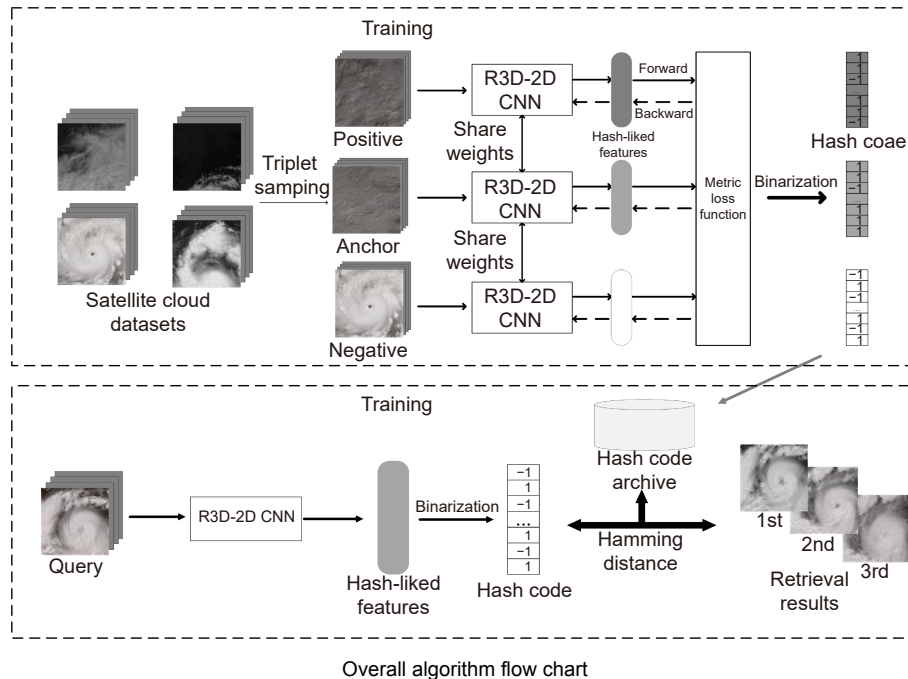


【通信作者】金炜 (1969-), 男, 博士, 教授, 研究方向: 数字图像处理。

E-mail: xyjw1969@126.com

Satellite cloud image retrieval based on deep metric learning

Jin Zhuzhang, Fang Xuyuan, Huang Yanhui, Yin Caoqian, Jin Wei*



Overview: Meteorological satellites can monitor weather phenomena of different scales from the air, and the satellite cloud images obtained by them play an important role in weather analysis and forecast. In recent years, with the development of meteorological satellite technology, the spatial and spectral resolution of satellite cloud images and the acquisition frequency of imaging spectrometer have been continuously improved. How to manage massive satellite cloud images and design an efficient cloud image retrieval system has become a difficult problem for meteorologists. However, the traditional cloud image retrieval methods are difficult to obtain ideal retrieval accuracy and retrieval efficiency. Motivated by the impressive success of the modern deep neural network (DNN) in learning the optimization features of specific tasks in an end-to-end fashion, a cloud image retrieval method based on deep metric learning is proposed in this paper. Firstly, a residual 3D-2D convolutional neural network was designed to extract spatial and spectral features of cloud images. Since the features extracted by the traditional classify-based deep network may have greater differences intra-classes than inter-classes, the triplet strategy is used to train the network, and the cloud images are mapped into the metric space according to the similarity between cloud images, so that the distance of similar cloud images in the embedded space is smaller than that of non-similar cloud images. In model training, the convergence performance of traditional triplet loss is improved and the precision of cloud image retrieval is increased by adding a constraint on the distance between positive sample pairs to the lossless triplet loss function. Finally, through hash learning, the cloud features in the metric space are transformed into hash codes, so as to ensure the retrieval accuracy and improve the retrieval efficiency. Experimental results show that the mean average precision (mAP) of the proposed algorithm is 75.14% and 80.14% for the southeast coastal cloud image dataset and the northern hemisphere cloud image dataset respectively, which is superior to other comparison methods.

Jin Z Z, Fang X Y, Huang Y H, et al. Satellite cloud image retrieval based on deep metric learning[J]. *Opto-Electron Eng*, 2022, 49(4): 210307; DOI: [10.12086/oee.2022.210307](https://doi.org/10.12086/oee.2022.210307)

Foundation item: National Natural Science Foundation of China (42071323) and Public Welfare Science and Technology Project of Ningbo (202002N3104).

Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, Zhejiang 315211, China

* E-mail: xyjw1969@126.com