

基于深度学习的电力线路航拍图像检索分析方法

刘主光¹, 杨 迁¹, 吕家辉²

(1. 浙江图盛输变电工程有限公司 温州科技分公司, 浙江 温州 325000;

2. 温州图盛控股集团有限公司, 浙江 温州 325000)

摘要: 电力行业的快速发展使得电力线路的巡检工作变得越来越重要; 传统巡检方法存在着效率低、精度不高等问题, 基于航拍图像的电力线路巡检已经成为重要的解决途径; 在传统图像检索框架的基础上, 创新提出一种基于深度学习的图像检索方法; 通过构建深度残差注意力网络哈希图像检索模型, 采用通道和空间注意力模块加强特征的细化, 自动识别并聚焦于图像中的关键区域; 此外, 通过基准样本的三元组损失函数进行端到端学习, 优化图像之间的距离度量, 从而实现相似图像的精确区分; 对比实验表明, 该系统能够有效地对电力线路航拍图像进行检索和分析, 为电力线路巡检工作提供强有力的支持。

关键词: 电力巡检; 图像检索; 分析系统; 深度学习; 注意力机制; 三元组损失函数

Retrieval and Analysis Method for Aerial Images of Power Lines Based on Deep Learning

LIU Zhuguang¹, YANG Qian¹, LÜ Jiahui²

(1. Wenzhou Science and Technology Branch, Zhejiang Tusheng Power Transmission and Transformation Engineering Co., Ltd., Wenzhou 325000, China;

2. Wenzhou Tusheng Holding Group Co., Ltd., Wenzhou 325000, China)

Abstract: With the rapid development of the power industry, the inspection of power lines has become increasingly important. Traditional inspection methods have the characteristics of low efficiency and low precision, so power line inspections based on aerial images have become an important solution. Based on the traditional image retrieval framework, an innovative image retrieval method based on deep learning is proposed. By constructing a deep residual attention network Hash image model, channel and spatial attention modules are used to enhance feature refinement, automatically identifying and focusing on key regions of the images. Additionally, the system utilizes a triplet loss function with benchmark sample for end-to-end learning, optimizing the distance metrics between the images to accurately distinguish similar images. Comparative experiments show that the system can effectively retrieve and analyze aerial images of power lines, providing strong support for power line inspections.

Keywords: power inspection; image retrieval; analysis system; deep learning; attention mechanism; triplet loss function

0 引言

随着电力需求的不断增长, 输电线路的安全保障至关重要。然而, 传统巡检方式受限于地形和天气条件, 效率低下且风险大。航拍图像技术通过无人机和航空摄影与测量技术, 可实现对输电线路高效、全方位的监

测, 提高巡检效率和安全性, 及时发现隐患与故障, 助力推动电力行业智能化发展。为此, 创新电力线路航拍图像检索方法具有迫切性和重要性, 对于提升电力线路管理水平、保障电力供应安全具有重要意义。

在无人机电力线路的巡检关键技术方面, 康泰安等^[1]提出了一种基于无人机航拍图像的电力线路故障检

收稿日期:2023-12-22; 修回日期:2024-02-02。

基金项目:电网企业开放基金项目(CF058804062021011)。

作者简介:刘主光(1974-),男,大学本科,高级工程师。

通讯作者:杨 迁(1993-),男,硕士,工程师。

引用格式:刘主光,杨 迁,吕家辉. 基于深度学习的电力线路航拍图像检索分析方法[J]. 计算机测量与控制, 2025, 33(3):295

- 301.

测方法，通过深度学习算法实现对故障点的自动识别和定位。赵振兵等人^[2]则进一步优化了无人机航拍系统，提高了图像质量和拍摄效率，为电力线路的全面监测提供了有力支持。电力线路航拍图像检索关键技术方面，杨慧等^[3]研究了基于深度学习的图像特征提取方法，提高了图像检索的准确性。周文罡等人^[4]则提出了一种基于哈希算法的图像匹配方法，有效降低了图像检索的计算复杂度。

随着大数据技术的发展，大量航拍图像数据的处理和分析成为了研究重点。刘传洋等^[5]提出了一种基于云计算的航拍图像处理框架，实现了对海量图像数据的高效处理和分析。然而，传统的电力线路航拍图像检索方法存在学习到的哈希码长度较短，对于复杂的场景学习能力不足，表达特征能力有限等问题。传统的三元组损失主要通过减小基准样本与正样本之间的间距，并增加基准样本与负样本之间的间距，但忽视了正样本与负样本之间的距离。因此，可能出现正样本与负样本之间的距离小于基准样本与正样本之间的距离，导致学习到的哈希码存在量化误差，从而影响检索精度。AletNet^[6]、VGGNet^[7]、GoogleNet^[8]、ResNet^[9]等网络证明，网络性能的提升关键在于网络设计，其深度和宽度能提高模型的表达能力，进而提高其性能。

为此，针对更高效的学习特征图表达并抑制离散过程中量化偏差的目标，论文提出了基于深度残差注意力网络的哈希图像检索方法（DRANH, deep residual attention network hashing）。三元组图像作为网络的输入，在深度残差网络的卷积层和全连接层之间嵌入注意力模块，从而使最后一个全连接层变成哈希层，通过对度量三元组损失学习到更好的相似性度量；注意力模块学习强调或抑制的内容和位置，并有效地改进中间特征，提升网络捕获有效图像信息的能力，从而提高特征

学习的能力。

1 DRANH 网络模型设计

DRANH 网络结构由卷积层、残差层、全连接层、注意力模块、哈希层和损失函数组成，用于生成有效的哈希码。残差网络也能够训练更广泛的网络结构，包括使用跳跃连接取代直接堆叠的层，或者直接把输入信息连接到后面的层，以维持信息的完整性，结构如图 1 所示。

图 1 中，给定输入图像在通道与空间的两个注意模块分别关注有意义的部分和最具信息量的部分。成对度量三元组损失继承了三元组损失的特点和优势，成对三元组度量损失能很好的部署在图像检索模型。

将三元组图像作为网络的输入，然后在 ResNet-50 层的最大池化层后面和全连接层中间插入注意力模块网络，将最后一个全连接层改为哈希层，最后通过对度量三元组损失学习到更好的相似性度量。采用了卷积注意力模型方法中的通道与空间注意力模型方法，在不修改特征图大小与通道数的情形下，将特征内嵌到网络的卷积层中，并通过端到端训练进行自适应的特征细化。注意力模型包括通道注意力模型和空间注意力模型，其架构如图 2 所示。

2 注意力机制

卷积操作将把通道和空间的信息混合在一起以获得图像特征。为此，采用卷积注意力模型来强调在通道和空间中重要的目标特征，再利用通道和空间注意模块，可以分别学习通道上有意义的部分和空间上最具信息量的部分。

通过学习强调并抑制信息来更高效地训练网络内的图像特征，或者通过更有效的架构同时使用空间与通道注意，并使用经验验证，使用这二者比只使用通道更有

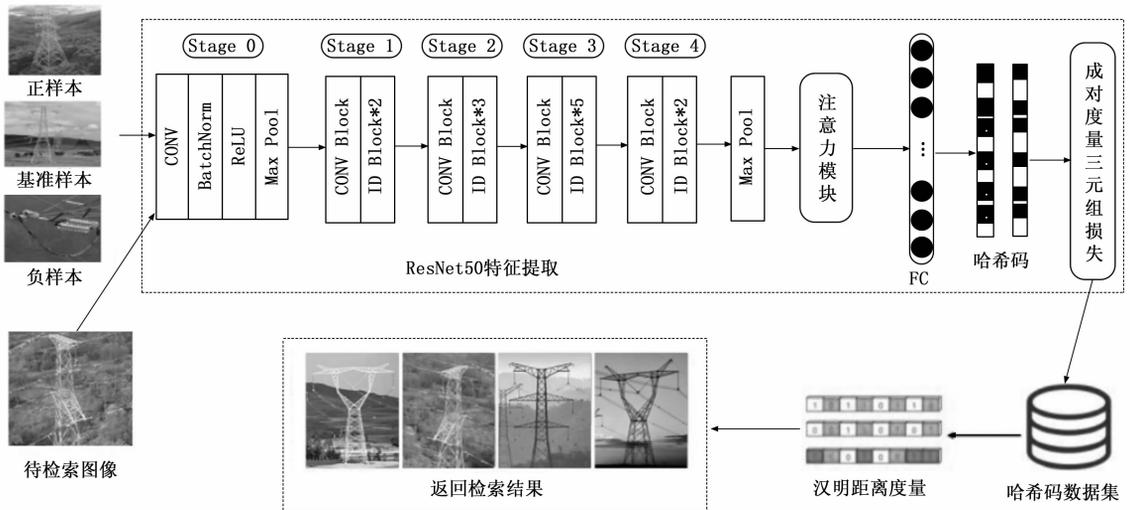


图 1 深度残差注意力网络哈希图像检索流程与结构图

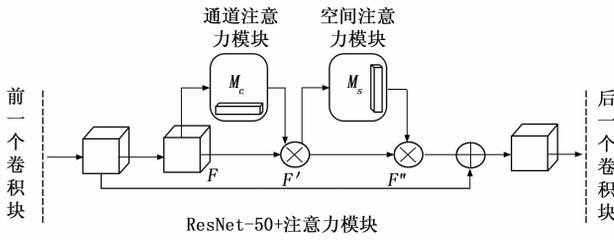


图 2 ResNet-50 与注意力模块结合图

效。深度残差网络的跳跃连接将输入连接到相应级别注意力模块, 通过处理图像特征, 能清楚识别最显著的对象和图像区域, 融合多级特征以改善检索性能。

输入给定的图像特征向量 $U \in R^{C \times H \times W}$, 通过运算导出二维矩阵空间注意力图 $M_s \in R^{1 \times H \times W}$ 和一维向量通道注意力图 $M_c \in R^{C \times H \times W}$, 注意力模块学习过程为:

$$\begin{aligned} U' &= M_c(U) \otimes U \\ U'' &= M_s(U) \otimes U' \end{aligned} \quad (1)$$

式中, \otimes 表示向量元素相对应的乘法, 一维向量通道注意力值沿二维矩阵空间维度复制。 U'' 为输出, 注意力图的运算流程如图 3 所示。

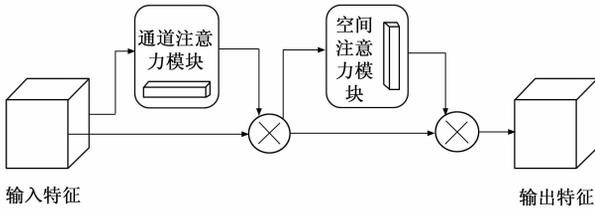


图 3 注意力模型图

2.1 通道注意力模块

高层特征图主要表示具有大感受野和图像的复杂特征, 需要通过学习全局图像信息来选择性地增强局部相关特征并抑制不相关特征。为此, 卷积层的最后一层添加通道注意模块。首先, 将特征图 $U \in R^{C \times H \times W}$ 实现聚合操作, 聚合操作利用聚合特征图在其空间维度 ($H \times W$), 生成通道注意力的一维向量特征:

$$q_c(U) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j) \quad (2)$$

式中, $U_c \in R^{H \times W}$ 为通道 c 的局部特征, 聚合函数 $q_c(U)$ 表示空间维度采用全局平均池化表示。

权重学习计算由全连接层实现, 函数 $F_1(q, w)$ 旨在完全捕获通道之间的依赖关系并自适应地生成通道权重映射 $v \in R^C$, 即:

$$v = F_1(q_1, w) = \sigma[g(q_1, w)] = \sigma[\omega_2 \delta(\omega_1 q)] \quad (3)$$

式中, $\omega_1 \in R^{K \times C}$, $\omega_2 \in R^{C \times K}$, K 为隐藏神经元个数, σ 是 *sigmoid* () 激活函数, δ 为 ReLU 激活函数, 用于在通道 c 处生成通道权重 $v_c \in (0, 1)$ 。利用全连接的隐藏层, 能够捕获通道内部的非线性交互。

计算通道特征图 U 将应用到权重, 通过将 U 中各个通道的特征值乘以不同的权重, 可以得到通道注意力模块的输出 U' :

$$U'_c = F(u_c, v_c) = u_c \times v_c \quad (4)$$

2.2 空间注意力模块

空间上的注意力将公平的对待不同通道的特征。由于低级特征图主要提取空间特征, 为此将空间注意模块添加到低级特征图。空间注意力模块自学习空间点的交互, 增强关键区域并抑制无关区域。首先, 将特征图 $U \in R^{C \times H \times W}$ 传递给聚合操作, 聚合操作生成空间描述符 $p \in R^{H \times W}$ 通过在其通道维度 C 中聚合特征图。生成空间特征的全局分布:

$$p_{h,w} = F(U_{h,w}) = \frac{1}{C} \sum_{i=1}^C u_{h,w}(i) \quad (5)$$

式中, $U_{h,w} \in R^C$ 为空间位置处 (h, w) 的局部特征, 聚合函数 F 表示通道维度采用全局平均池化表示。

权重学习计算由卷积层实现, $F_1(p, f)$ 函数主要充分捕捉空间相关性并自适应生成空间权重图 $w \in R^{H \times W}$, 计算公式如下:

$$w = F_1(p_1, f) = \sigma[g(p_1, f)] = \sigma[f_2 \delta(f_1 p)] \quad (6)$$

式中, f_1 为 3×3 的卷积块, f_2 也是 3×3 卷积块, 记为 $\text{Conv}(3 \times 3, 1)$ 。 m 指隐藏特征图的通道数。 σ 是一个 *sigmoid* () 激活函数, δ 为 ReLU 激活函数, 用于在 (h, w) 位置生成空间权重由 $t_{h,w} \in (0, 1)$ 。本质上以原始空间描述符为输入的卷积运算, 捕捉非线性的空间关系。

计算得到的权重应用于特征图 U 。通过空间重校准, 将 U 中不同位置的特征值乘以不同的权重, 生成注意力模块的输出 U' :

$$u'_{h,w} = u_{h,w} \times t_{h,w} \quad (7)$$

2.3 对比距离损失

基准样本的三元组损失函数通过学习样本类中心, 将其用作基准来形成三元组, 增加代表基准样本和负样本之间各种特征向量的距离。图像检索网络模型的设计主要由 softmax 交叉熵损失和三元组损失组成。 N 表示调整参数前所选的样本的数量, $p(k)$ 表示为第 k 类预测, 而必表示真值类, 则交叉熵损失 L_{id} 计算为:

$$L_{id} = - \sum_{k=1}^i q(k) \log[p(k)] \quad \begin{cases} q(k) = 0, y \neq k \\ q(k) = 0, y = k \end{cases} \quad (8)$$

标签平滑将用于防止的过拟合, 表示为:

$$q(k) = \begin{cases} \frac{1 - (N - 1)\lambda}{N}, k = y \\ \lambda, \text{otherwise} \end{cases} \quad (9)$$

式中, λ 是一个常数, 用于减小模型的置信度。如果基准样本、正样本和负样本的全局特征分别表示为 f_a 、 f_p 和 f_n , 则 m 为阈值, 三元组损失通常表示为:

$$L_{tri} = - \sum_{i=1}^p \sum_{a=1}^K \left[\begin{array}{c} m + \max_{p=1 \dots K} \|f_a^i - f_p^i\|_2 \\ - \min_{j=1 \dots K} \|f_a^i - f_j^i\|_2 \\ i \neq j \end{array} \right]_+ \quad (10)$$

一个小批次样本中正负样本对有 P 对，每对中有 K 张图片， $[\cdot]_+$ 表示当 $[\cdot]$ 中的值小于等于 0 时，损失值为 0，当大于 0 时，取该值作为损失。三元组损失使用欧式距离度量来进行样本挖掘。

传统的三元组损失使用单个欧式距离作为特征空间的度量，这已被广泛证明是一种有效的度量学习损失。然而，从单个图像中测量特征的方法相对有偏差。因此，提出了成对度量三元组 (Pairwise-Triplet) 损失，引入了成对距离度量，利用欧式距离和余弦距离的优势，互补提高模型的性能。从成对距离度量的角度采取不同的观点，以稳定和通用的方式学习判别特征以进行图像检索。在 Pairwise-Triplet 损失中引入余弦距离减小偏差，与每个点的绝对位置坐标相关的欧式距离则根据相似度矩阵提取样本对，并计算 Pairwise-Triplet 损失的结果。余弦距离测量了二种特征向量间的余弦角间距，从而测量它们间的相似性，以降低哈希码的测量误差。如图 4 所示成对度量三元组损失的示意图。

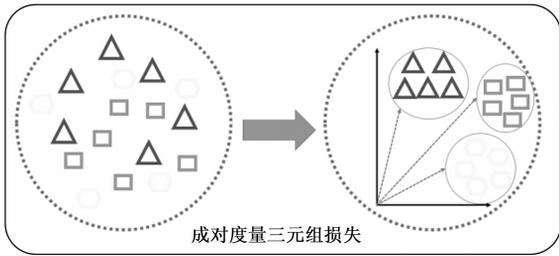


图 4 成对度量三元组损失示意图

图 4 中，Pairwise-Triplet 损失引入另一个度量后，模型将从两个角度进行样本学习，以削弱三元组损失只有一个度量的偏差。余弦距离挖掘和区分不同于欧式距离度量的信息，余弦夹角则用来区分具有相同欧式距离的样本。由于不同指标提供的样本特征信息不同，这两种信息在一定程度上互补学习，有助于模型扩大样本的相似性特征，从而有效提高模型的性能。

若骨干网输出特征张量表示为 $\mathbf{X} \in R^{b \times c}$ ， \mathbf{X} 可以表示为：

$$\mathbf{X} = \begin{bmatrix} f^1 \\ f^2 \\ \vdots \\ f^b \end{bmatrix} = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_c^1 \\ x_1^2 & x_2^2 & \dots & x_c^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^b & x_2^b & \dots & x_c^b \end{bmatrix} \quad (11)$$

式中， b 和 c 分别是批量大小和特征向量的维度。而 x_c^1 ， x_1^b 分别是第一个特征向量的最后一个特征元素和第 b 个特征向量的第一个特征元素。

余弦距离相似度矩阵 $dist_c$ 表示为：

$$dist_c = \frac{\mathbf{X} \cdot \mathbf{X}^T}{\|\mathbf{X}\| \|\mathbf{X}^T\|} \quad (12)$$

对于小批次样本中的第 i 个特征向量 $\|f^i\|$ ($1 \leq i \leq b$)，

$$\|f^i\|_2 = \sqrt{(x_1^i)^2 + (x_2^i)^2 + \dots + (x_c^i)^2} \quad (13)$$

其中： $dist_c$ 以矩阵形式计算了小批次样本中成对特征向量之间的相似度。计算余弦距离样本损失 L_c ，计算结合余弦度量矩阵：

$$dist_c = clamp(\mathbf{X} \cdot \mathbf{X}^T, -1 - \zeta, 1 + \zeta) \quad (14)$$

式中， $clamp()$ 函数用于将目标值截断到一定的区间，防止出现异常值， ζ 是一个常量，推导出余弦样本对矩阵 h_p 、 h_n 和标签 y 。余弦距离的损失：

$$L_c = max[0, -y \cdot (h_n - h_p) + m] \quad (15)$$

计算欧式距离矩阵，进而推导出欧式距离样本对矩阵 h_p 、 h_n 和标签 j ，欧式距离的样本损失 L_e ：

$$L_e = max[0, -y \cdot (h_n - h_p) + m] \quad (16)$$

成对度量三元组损失使用余弦角推导出余弦相似度矩阵。相似度矩阵 $dist_c$ 将余弦值转换为余弦角为：

$$dist_c = a \cos\left(\frac{\mathbf{X} \cdot \mathbf{X}^T}{\|\mathbf{X}\| \|\mathbf{X}^T\|}\right) \quad (17)$$

余弦角在度量学习将值映射到更广泛的范围，通过测量它们之间的余弦角来量化小批次样本中每两个特征向量对之间的相似度。利用二个相似度矩阵，可以使用前面公式推导出正负样本对，然后按照相似度距离对样本对进行排序，选取相同标签中相似度最少的正样本，并选取不同标签中相似度最高的负样本。三元组损失通过对网络参数的深度学习，既减少了正样本之间的距离，同时也增加了负样本间的距离。

3 实验与分析

3.1 实验设置

在本节中，先在两种基准图像数据集 MBP^[10] 和 CPLID^[11] 中进行实验，接着进行了通道与空间注意力的排列顺序的消融实验，随后再验证中加入了成对度量三元组损失的消融实验，最后又与 8 种深度哈希图像检索方法进行了全面对比，对所提出的方法进一步分析。为了全面评估本文的方法以及所有对比方法的检索准确度，选择了普遍使用的平均精度均值 mAP 和准确率—召回率 P-R 曲线图作为实验结果的主要衡量标准，并使用汉明距离计算图像间的相似度。

实验在 Windows10 操作系统，显卡为 GeForceRT-X2080TiGPU，内存为 32 GB 的计算机上完成，使用 Pytorch 框架实现的，ResNet-50 为主干网络模型，在实验训练过程中，使用 mini-batch 优化器，将批次大小设置为 256，动量设置为 0.9，重量衰减设置为 $5e-4$ ，学习率为 0.001，总共训练了 150 个 epoch，参数为哈

希码的位数, 取值为 12、24、32 和 48。

MBP 数据集主要是包含同一个目标拍摄的九段视频, 并进行分帧处理。包含 2 630 张图片, 分辨率是 5 616×3 744。本实验的训练集有 1 840 张图片, 余下的约一万张图片则为测试集。

CPLID 包含 600 张真实世界的和粘贴在不同背景的图像。有缺陷的绝缘子图像有 248 张, 图像尺寸为 1 152×864。在实验中, 使用在 21 个最常出现的语义标记中随机选取 100 张图像用作测试集, 而其余图像用作训练集。

论文将以基准数据集 MBP 和 CPLID 为基线开展图像检索实验, 用于评估深度残差注意力网络模型。为检验模型的有效性, 和九种最经典和最先进的基于深度学习的哈希方法加以对比, 包括卷积神经网络哈希 (CNNH, convolutional neural network hashing)^[12], 深度神经网络哈希 (DNNH, deep neural network hashing)^[13], 深度费希尔哈希 (DFH, deep fisher hashing)^[14], 改进深度哈希网络 (IDHN, improved deep hashing network)^[15], 深度极化网络 (DPN, deep polarized network)^[16], 深度平衡离散哈希 (DBDH, deep balanced discrete hashing)^[17], 哈希网络 (HashNet, hash network)^[18], 深度成对监督哈希 (DPSH, deep pairwise supervised hashing)^[19], 深度三元组监督哈希 (DTSH, deep triplet supervised hashing)^[20]。

1) CNNH: 哈希码学会首先从成对的标签开始学习哈希码, 而后试图利用哈希码在图形像素中掌握哈希函数和特性。

2) DNNH: 深度神经网络哈希可以通过深度神经网络产生卷积层的子网络, 映射成二进制哈希码, 形成有效的中间图像特征, 使用三元组损失函数对比相似性。

3) DFH: 采用线性判别分析最大化类之间的二进制距离, 同时最小化同一类内图像的二进制距离。

4) IDHN: 改进的深度哈希方法来增强多标签图像检索的能力, 使用一对多标签图像之间的细粒度相似性以进行哈希学习。

5) DPN: 用于学习哈希的新型深度极化网络, 最小化偏振损失相当于同时最小化内部方差和最大化类间方差的汉明距离。

6) DBDH: 深度平衡离散哈希使用带直通估计器的离散梯度传播, 通过监督信息指导和深度特征学习并离散编码, 组合成对损失和平衡控制项计算离散值。

7) HashNet: HashNet 是一种新颖的深度学习架构, 具有收敛保证的连续方法进行哈希编码, 在非平滑二元激活的深度网络优化梯度问题。

8) DPSH: 深度成对监督哈希, 通过最大化包含

成对标签的应用程序并且同时执行特征学习和哈希码学习。

9) DTSH: 基于三元组标签的深度哈希方法, 以三元组标签的形式提供监督信息, 并通过深度神经网络同步学习图像特征和哈希码, 旨在最大化标签的可能性。

3.2 消融实验

选择以 ResNet-50 为基础架构, 在 MBP 数据集上进行实验。模块设计过程分为两个部分, 首先验证通道注意力和空间注意力的有效方法及排序, 之后是验证加入成对度量量化三元组损失函数的性能提升。

1) 注意力模块排序:

实验比较了通道与空间注意力模块之间的三个不同排列方式: 先通道后空间、先空间后通道、通道和空间模块并行。由于各个模块都有着不同的用途, 分别关注最有意义的部分和最具信息量的部分, 因此排列顺序将会影响模型整体性能, 因此需要找出最优的排列顺序。由形成的一维通道注意力图和二维空间注意力图组合, 形成了三维空间的注意力图。

首先输入图像, 通道注意力和空间注意力模块能够很好的生成三维注意力图。实验结果如表 1 所示。

表 1 比较不同的排序注意力方法 %

	Top-1Acc	Top-5Acc
ResNet-50	75.34	79.63
ResNet-50+通道+空间	81.13	91.38
ResNet-50+空间+通道	80.39	90.35
ResNet-50+通道空间平行	79.16	90.14

表 1 表明并行排列的性能略低于顺序排列。ResNet-50 表示没有加入注意力机制时的检索准确率, 检索准确率达到 75.34%, 相比加入注意力机制准确率降低 5.79%, 表明注意力机制能增强模型检索准确性和有效性。对于顺序排序, 实验结果显示, 通道优先略好于空间优先。通道注意力与空间注意模块的组合, 对使用这两种注意力模块至关重要, 而最佳组合策略 (即先通道后空间的顺序) 进一步提高了准确性, 最终模块如图 3 所示, 实现了 81.13% 的准确率。

2) 成对度量三元组损失:

实验中, 在 ResNet-50 加入注意力模型 (先通道后空间顺序), 在 MBP 数据集上进行了实验, 并以 ResNet-50 的网络模型为实验基线, 通过对比改进三元组损失后的性能。成对度量三元组损失通过结欧式距离度量和余弦距离度量, 学习全局特征来改善度量学习损失。通过选择 ResNet-50 作为基线来测试加入损失函数后的图像检索性能, 比较加入三元组损失和成对度量三元组损失的性能。结果如表 2 所示。

表 2 比较不同损失函数的性能 %

	Top-1 Acc	Top-5 Acc
ResNet-50	81.13	91.38
ResNet-50+三元组损失	82.43	92.58
ResNet-50+成对度量三元组损失	83.67	93.86

在基准数据集 MBP 上, 具有成对度量三元组损失的 ResNet-50 在 Top-1 准确性中超过了 ResNet-50 基线 2.54% 的模型, 比加入三元组损失准确性超过了 1.24%。结果性能表明, 成对度量三元组损失在不调整任何训练配置的情况下提高三元组损失在主干网络模型上的性能, 这证明了成对度量三元组损失继承了三元组损失良好的泛化能力。成对度量三元组损失通过探索相似性信息来增强样本的检索精度, 有效提高基线的性能, 并且与骨干网络模型兼容, 表明成对度量三元组损失在正负样本对挖掘中的能力更具针对性, 引入成对度量可以挖掘出单个度量忽略的信息, 提高了具有相同标签的相似样本的精度。

3.3 实验结果与分析

本文将表示了深度残差注意力网络的哈希图像检索方法 (DRANH) 和九种不同深度哈希图像检索方法作

为对比, 在 MBP 和 CPLID 两个基准数据集上进行实验, 得出了不同长度 (12 位、24 位、32 位和 48 位) 的哈希码图像检索性能结果, 如表 1 所示。可以看到, 与具有不同长度的哈希码与其他方法相比, DRANH 的检索精度具有一定的提高。如表 3 与图 5 所示, 本文提出的 DRANH 方法的性能通常随着哈希码长度的增加而提高。由于哈希码长度的增加, 学习的图像特征更丰富, 进而提高检索精度。

MBP 数据集的实验结果表明, 本文提出的 DRANH 方法的性能显著高于其他所有方法, DRANH 使用残差注意力网络提取图像特征信息, 哈希码的检索性能在不同长度上达到了 79.1%, 80.7%, 82.1% 和 83.6%, 比使用端到端网络训练的 DNNH 方法相比, 检索性能分别提升了 23.9%、24.1%、26.3% 和 25.5%, 与使用三元组损失函数的 DTSH 方法相比, 检索性能提升了 1.9%、1.8%、1.8% 和 2.5%, 性能略有提升。

论文对大规模图像数据集进行了实验, CPLID 数据集规定假如两个图像至少共用一个标签, 则认为它们属于同一类别。表 3 实验结果表明, 所提出的 DRANH 方法在大规模图像集上的检索性能在不同哈希码上分别达到了 79.5%、82.6%、84.3% 和 85.4%, 与最经典最

表 3 在 2 个数据集上不同哈希码长度的 mAP 值

方法	MBP				CPLID			
	12-bit	24-bit	32-bit	48-bit	12-bit	24-bit	32-bit	48-bit
CNNH	0.439	0.517	0.512	0.523	0.611	0.618	0.625	0.618
DNNH	0.552	0.566	0.558	0.581	0.674	0.697	0.713	0.715
DFH	0.752	0.773	0.791	0.802	0.775	0.816	0.825	0.844
IDNH	0.744	0.746	0.768	0.781	0.784	0.793	0.801	0.804
DPN	0.755	0.759	0.789	0.769	0.762	0.793	0.809	0.827
DBDH	0.767	0.790	0.779	0.782	0.802	0.832	0.836	0.841
HashNet	0.537	0.766	0.779	0.786	0.748	0.793	0.820	0.841
DPSH	0.786	0.783	0.796	0.776	0.799	0.823	0.833	0.837
DTSH	0.772	0.789	0.803	0.811	0.801	0.825	0.840	0.847
DRANH	0.791	0.807	0.821	0.836	0.795	0.826	0.843	0.854

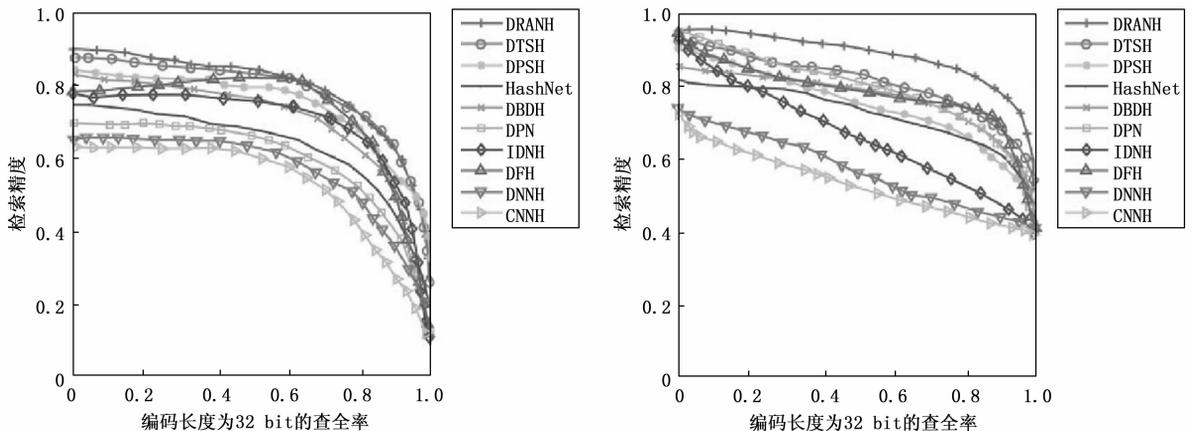


图 5 两个数据集在 32 bits 编码下的查准率、查全率和 P-R 曲线图

先进的深度哈希方法相比,DRANH方法的检索性能略有提高。与使用成对损失和平衡控制项计算哈希值的深度平衡离散哈希(DBDH)相比,检索性能提升了 -0.7% , -0.6% , 0.7% 和 1.3% 。与使用三元组损失的监督哈希(DTSH)相比,检索性能提升了 -0.6% , 0.1% , 0.3% 和 0.7% 。

在MBP和CPLID基准数据集上的实验结果表明,本文的DRANH方法普遍高于当前的深度哈希方法,包括性能较好的深度平衡离散哈希(DBDH)和深度三元组监督哈希(DTSH),实现了最佳的检索性能。表明在残差网络中加入了注意力模块,进行特征学习过程中削弱了无效信息的干扰,以及加入了成对度量三元组损失以优化三元组损失,减小了生成的哈希码的量化误差。

4 结束语

论文提出了基于深度残差注意力网络的图像检索方法,通过通道和空间注意力模块应用加强特征的细化,改进损失函数提高哈希码的学习能力,实现了强大的性能提升。注意力模块学习了强调或抑制的内容和位置,并有效地改进了中间特征。在哈希层后引入成对三元组度量损失,结合欧式距离和余弦距离的度量方法,对传统三元组损失进行改进。通过这种方式,模型能够更好地学习全局特征,从而优化度量学习损失,提高检索性能。MBP和CPLID基准数据集上的实验,证明了先通道后空间注意力模块的有效性和加入成对度量三元组损失的性能提高;与最先进的图像检索算法比较实验,结果表明所提方法能有效的强调有用特征来更好的训练网络,有助于扩大样本间的相似性特征,从而提高模型的性能,图像检索性能得到了改善。

参考文献:

- [1] 康泰安, 王宝莉, 刘少航, 等. 输电线路金具及其缺陷深度学习检测方法综述[J]. 电力信息与通信技术, 2022, 20(11): 1-12.
- [2] 赵振兵, 齐鸿雨, 聂礼强. 基于深度学习的输电线路视觉检测研究综述[J]. 广东电力, 2019, 32(9): 11-23.
- [3] 杨慧, 施水才. 基于内容的图像检索技术研究综述[J]. 软件导刊, 2023, 22(4): 229-244.
- [4] 周文罡, 李厚强, 田奇. 图像检索技术研究进展[J]. 南京信息工程大学学报(自然科学版), 2017, 9(6): 613-634.
- [5] 刘传洋, 吴一全, 刘景景. 无人机航拍图像中电力线检测方法研究进展[J]. 中国图象图形学报, 2023, 28(10): 3025-3048.
- [6] GOLUGURI R V N, K D S, GOGULA D S, et al. AT-RLNet: A deep learning model for enhanced classification of *Oryza sativa* pathologies[J]. Traitement du Signal, 2023, 40(4):
- [7] SHENG F, LIPING Z, HAIYANS, et al. One-dimensional VGGNet for high-dimensional data[J]. Applied Soft Computing Journal, 2023, 135
- [8] SEMENYUK V, KURMASHEV I, LUPIDI A, et al. Developing the GoogleNet neural network for the detection and recognition of unmanned aerial vehicles in the data Fusion System[J]. Eastern-European Journal of Enterprise Technologies, 2023, 2(9): 16-25.
- [9] SAJEDI H, MOHAMMADIPANAH F, RAHIMI H A S. Actinobacterial strains recognition by machine learning methods[J]. Multimedia Tools and Applications, 2019, 78(14): 20285-20307.
- [10] MICHAT TOMASZEWSKI, Bogdan Ruszczak, Pawet Michalski (2018), The collection of images of an insulator taken outdoors in varying lighting conditions with additional laser spots, Data in Brief. <https://github.com/InsulatorData/InsulatorDataSet>.
- [11] XIA R, PANY, LAI H, et al. Supervised hashing for image retrieval via image representation learning[C] // Twenty-eighth AAAI conference on artificial intelligence, 2014.
- [12] LAI H, PANY, LIU Y, et al. Simultaneous feature learning and hash coding with deep neural networks[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3270-3278.
- [13] LI Y, PEI W, VANGEMERT J. Push for quantization: Deep fisher hashing[J]. ArXiv Preprint ArXiv: 1909.00206, 2019.
- [14] ZHANG Z, ZOU Q, LIN Y, et al. Improved deep hashing with soft pairwise similarity for multi-label image retrieval[J]. IEEE Transactions on Multimedia, 2019, 22(2): 540-553.
- [15] FANL, NG K W, JU C, et al. Deep polarized network for supervised learning of accurate binary hashing codes[C] // IJCAI. 2020: 825-831.
- [16] ZHENG X, ZHANG Y, LU X. Deep balanced discrete hashing for image retrieval[J]. Neurocomputing, 2020, 403: 224-236.
- [17] CAO Z, LONG M, WANG J, et al. Hashnet: Deep learning to hash by continuation[C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 5608-5617.
- [18] LI W J, WANG S, KANG W C. Feature learning based deep supervised hashing with pairwise labels[J]. ArXiv Preprint ArXiv: 1511.03855, 2015.
- [19] WANG X, SHI Y, KITANI K M. Deep supervised hashing with triplet labels[C] // Asian Conference on Computer Vision, Springer, Cham, 2016: 70-84.