

# Asymmetric Transfer Hashing With Adaptive Bipartite Graph Learning

## 摘要

由于高效的检索速度和低存储消耗, 哈希学习已广泛应用于视觉检索任务中。然而, 已知的哈希方法假设查询和检索样本位于同一域内的同构特征空间中。因此, 它们不能直接应用于异构跨域检索。在本文中, 提出了广义图像迁移检索 (GITR) 问题, 该问题遇到两个关键瓶颈: 1) 查询和检索样本可能来自不同的域, 导致不可避免的域分布差距; 2) 两者的特征域可能是异构的或非对齐的, 从而带来额外的特征隔阂。为了解决 GITR 问题, 提出了一种非对称迁移哈希 (ATH) 框架及其无监督/半监督/监督实现。具体来说, ATH 通过两个非对称哈希函数之间的差异来表征域分布差距, 并借助在跨域数据上构建的新型自适应二部图来最小化特征差距。通过联合优化非对称哈希函数和二部图, 不仅可以实现知识转移, 还可以避免特征对齐带来的信息丢失。同时, 为了减轻负迁移, 通过使用亲和图来保留单域数据的内在几何结构。在不同 GITR 子任务下对单域和跨域基准进行的广泛实验表明, 与最先进的哈希方法相比, 所提出的 ATH 方法具有优越性。

**关键词:** 自适应二分图: 非对称哈希: 哈希学习: 迁移学习

## 1 引言

由于检索速度快和内存占用少, 哈希学习在过去的几十年里在不同的视觉任务中引起了广泛的关注。哈希方法旨在将数据编码为紧凑的二进制代码, 以便基于硬件级异或运算进行快速检索。最近, 人们从不同的角度提出了大量的哈希方法。然而, 已知的哈希方法主要依赖于两个假设: 1) 需要足够的训练样本来学习判别性哈希函数; 2) 查询和检索样本来自具有相同数据分布的同一域。在现实场景中, 这些假设常常很容易被违反。具体来说, 感兴趣领域中的可用数据可能不足以进行训练, 从而导致数据稀疏问题。此外, 查询域和检索域之间可能不可避免地存在数据分布差异。为了解决这些缺点, 一个可行的解决方案是将哈希方法与迁移学习 (TL) 相结合。TL 旨在通过迁移不同但相关的源域中包含的知识来增强目标求解器/学习器在目标域上的性能。因此, 可以减轻目标求解器构建对足够目标域数据的依赖。

受 TL 的启发, 最近提出了一些传输哈希 (TH) 方法。第一个 TH 方法称为 TH 与特权信息 (THPI) [1], 结合哈希和 TL 方法来解决数据稀疏问题。与 THPI 不同, 最优投影引导 TH (GTH) [2] 寻求目标域和源域的哈希函数的最大似然估计解决方案。新开发的概率加权紧凑特征学习 (PWCF) 提供域间相关性指导以提高跨域检索精度, 判别哈希学习 (DHLing) [3] 构建域不变存储库以实现跨域对齐。然而, 上述方法都是同构 TH 方法, 不能直接应用于异构域。具体来说, 这些方法假设目标域和源域共享具有相同特征维度的相同特征空间。

这个假设可能并不总是成立，因为两者的特征域可能是异构的或者非对齐的。例如，查询和检索样本可能具有不同的分辨率和大小，由不同类型和长度的手工特征表示（例如，颜色直方图和 SIFT），或者由独特的深度神经网络提取（例如，VGG [4] 和 ResNet [5]）。在这些情况下，已知的 TH 方法需要额外的特征预处理（例如图像缩放和降维）来进行特征对齐，这可能会在一定程度上导致信息丢失。

在本文中，首先总结了一种更具挑战性和实用性的迁移检索任务，称为广义图像迁移检索（GITR），其中查询和检索图像可能来自不同的领域，并且它们的特征可能是同质的或异构的。图 1 给出了所提出的 GITR 的说明，其中包含四个不同的子任务，包括异构跨域检索（HeCDR）、同质跨域检索（HoCDR）、异构单域检索（HeSDR）和同质单域检索（何特别报告员）。在对查询和检索集的特征空间和域没有任何限制的情况下，本文旨在开发一个通用框架来解决复杂的 GITR 问题。具体来说，给定在 D1 域的 F1 特征空间中观察到的查询图像，我们的目标是检索 D2 域的 F2 特征空间中表示的最相似图像，无论这两个域和两个特征空间是否相同。为了实现这一目标，提出了一个非对称 TH(ATH) 框架及其无监督/半监督/监督实现。所提出的 ATH 包含三个主要组成部分：1) 非对称哈希学习，它通过两个非对称哈希函数的差异来表征域分布差距；2) 跨域信息交互，优化自适应二分图，用于跨域数据的知识传递和信息融合；3) 域结构保留，通过基于域亲和图维护单域数据的内在几何结构来减轻负迁移。本文的主要贡献可总结如下。

1) 将当前的学习哈希问题扩展到更普遍和更具挑战性的情况，称为 GITR，它对查询和检索集的域和特征空间没有假设或限制。

2) 为了解决 GITR 问题，提出了一个 ATH 框架及其无监督/半监督/监督实现。具体来说，ATH 找到两个非对称哈希函数将样本编码为有判别性的哈希码，并学习自适应二分图来表征跨域样本之间的相似性。

3) ATH 是第一个在统一学习目标中联合优化非对称哈希函数和自适应二分图的异构 TH 框架，同时解决了 GITR 中的两个关键问题（见图 2）。

4) 在单域和跨域基准测试上进行的大量实验表明，针对特殊特征空间和域定制的已知哈希方法无法达到令人满意的性能，而 ATH 框架可以在不同场景下良好工作。

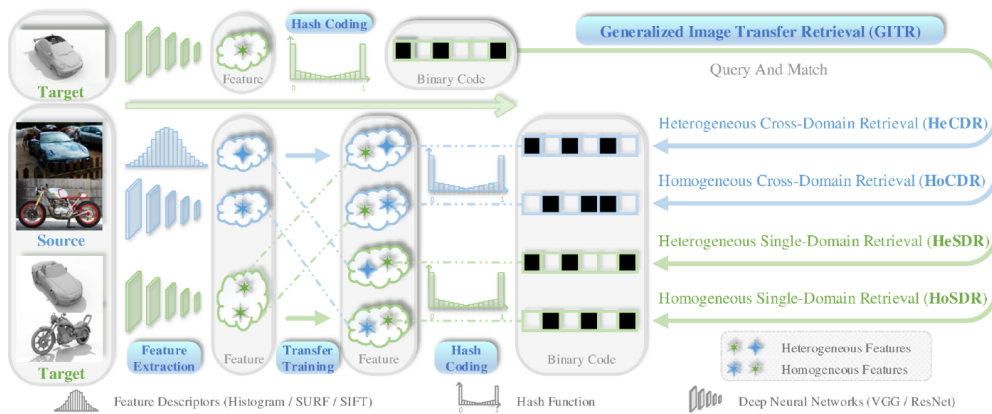


图 1. 所提出的 GITR 问题的说明

## 2 相关工作

### 2.1 泛化的图像迁移检索

假设我们有来自两个域的  $\mathcal{D} = \{\mathcal{D}_s, \mathcal{D}_t\}$  的  $n = n_s + n_t$  数据样本, 其中源域  $\mathcal{D}_s$  和目标域  $\mathcal{D}_t$  分别包含  $n_s$  和  $n_t$  样本。对于每个域  $\mathcal{D}_k, k \in \{s, t\}$ , 我们有  $\mathcal{D}_k = \{X_k, Y_k\} = \{(x_{k,i}, y_{k,i})\}_{i=1}^{n_k}$ , 其中  $x_{k,i} \in \mathbb{R}^{d_k}$  表示  $X_k \in \mathbb{R}^{r \times n_k}$  的第  $i$  个样本维度为  $d_k$ ,  $y_{k,i} \in \mathbb{R}^c$  是样本  $x_{k,i}$  对应的 one-hot 编码标签向量。为了简单起见, 我们假设源域和目标域中都有  $c$  个类。数据  $X_k$  的哈希码用  $B_k \in \mathbb{R}^{r \times n_k}$  表示, 其中  $r$  是二进制码的长度。此外,  $\mathbf{I}, \mathbf{1}, \mathbf{0}$  分别表示大小兼容的单位矩阵、全 1 的列向量和全 0 的列向量。

在 GITER 问题中, ATH 算法不要求查询和检索样本来自同一域。此外, 也不限制两个域的特征空间是同构的。因此, 先前 TH 方法中存在的假设  $d_s = d_t$  可能会被违反。如图 1 所示, HeSDR 和 HoSDR 子任务存在域一致性, HeCDR 和 HoCDR 子任务存在域多样性。此外, HoSDR 和 HoCDR 子任务存在特征同构性, HeSDR 和 HeCDR 子任务存在特征异构性。

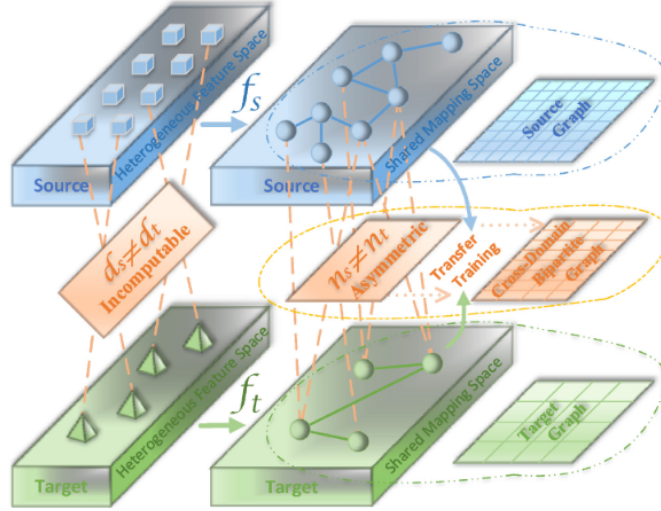


图 2. 从技术上讲, GITER 的面临的两个困难

### 2.2 同构迁移哈希

第一种 TH 方法 THPI 将来自源域的特权信息合并到目标域中以辅助哈希学习。其目标函数可以表述如下:

$$\begin{aligned} \min_{\mathbf{B}_t, \mathbf{A}_l, \mathbf{A}_{sc}} \quad & \mathbf{E}^2 + \lambda_1 \|\mathbf{E} - \mathbf{A}_{sc}^T \mathbf{X}_{sc}\|^2 \\ \text{s.t.} \quad & \mathbf{A}_l^T \mathbf{A}_l = \mathbf{I}, \quad \mathbf{A}_{sc}^T \mathbf{A}_{sc} = \mathbf{I} \end{aligned} \quad (1)$$

其中  $\mathbf{E} = \mathbf{B}_t - \mathbf{A}_l^T \mathbf{X}_t$ ,  $\mathbf{X}_{sc} \in \mathbb{R}^{d_s \times n_t}$  是从源域  $\mathcal{D}_s$  中选择出来的  $n_t$  个样本集合,  $\mathbf{A}_l \in \mathbb{R}^{d_t \times r}$  和  $\mathbf{A}_{sc} \in \mathbb{R}^{d_s \times r}$  分别是  $X_t$  和  $X_{sc}$  的投影矩阵。

可以看出, THPI 利用源域的数据信息来近似目标域的量化误差。由于  $n_t \neq n_s$ , THPI 只能利用  $n_s$  个源域样本与  $n_t$  个目标域样本进行信息交互, 这可能会削弱目标学习器的传输能力。此外, 如何从整个源域中选取  $n_t$  个样本仍有待确定。

与 THPI 不同, GTH 不需要从源域中选择一部分数据进行跨域信息交互。其目标可以表述如下:

$$\min_{\mathbf{B}_t, \mathbf{B}_s, \mathbf{A}_t, \mathbf{A}_s} \left\| \mathbf{M}^{\frac{1}{2}} \odot (\mathbf{A}_t - \mathbf{A}_s) \right\|^2 + \lambda_1 \left\| \mathbf{B}_t - \mathbf{A}_t^T \mathbf{X}_t \right\|^2 + \lambda_2 \left\| \mathbf{B}_s - \mathbf{A}_s^T \mathbf{X}_s \right\|^2, \text{s.t. } \mathbf{A}_t^T \mathbf{A}_t = \mathbf{I}, \mathbf{A}_s^T \mathbf{A}_s = \mathbf{I} \quad (2)$$

其中  $\odot$  表示逐元素乘积,  $\mathbf{M}$  是从最大似然估计导出的误差矩阵,  $\mathbf{A}_s \in \mathbb{R}^{d_s \times r}$  和  $\mathbf{A}_t \in \mathbb{R}^{d_t \times r}$  分别是域  $\mathcal{D}_s$  和  $\mathcal{D}_t$  的哈希投影。GTH 假设目标域和源域之间图像的相似性可以通过两个域的哈希投影之间的差异来反向表征。换句话说, 哈希投影  $\mathbf{M}^{(1/2)} \odot (\mathbf{A}_t - \mathbf{A}_s)$  之间的加权差异应最小化, 以确保来自目标域和源域的相似图像可以转换为相似的哈希码。这样, 源域中的全部知识都可以用于目标学习者。

PWCF 不是最小化哈希投影之间的差异, 而是设计了焦点三重态损失来促进跨域相关性并获得良好的传输性能。此外, DHLing [3] 进一步考虑了注释源图像不足的问题, 并提供了半监督方法。然而, 所有这些 TH 方法都假设目标域和源域共享具有对齐特征的同质特征空间 (即  $d_s = d_t = d$ )。由于检索和查询样本的特征维度可能不对齐, 因此该假设可能并不总是成立。例如, 查询和检索样本可能具有不同的分辨率和大小, 由不同长度的手工特征表示, 或者由不同的深度神经网络提取。在这些情况下, 这些方法不能直接应用于 HeSDR 和 HeCDR 子任务 (其中  $d_s = d_t$ ), 这促使我们开发一种新颖的异构 TH 方法来缓解这个问题。

## 2.3 非对称哈希

早期对非对称哈希的探索采用非对称距离度量来保持相似性, 其中训练集和测试集的哈希码是从相同的哈希函数生成的。内沙布尔等人。[6] 已经证明, 对查询和检索集使用不同的哈希函数可以用更短的代码长度更好地近似目标相似度。在这项工作之后, 最近的非对称哈希方法倾向于关注功能空间, 其中训练集和测试集之间的分布差距以两个不同哈希函数的差异为特征。一般来说, 这些方法采用用于相似性保存的相似性差异函数。

$$\sum_i \sum_j \|f_1(\mathbf{x}_i) f_2(\mathbf{x}_j) - r \mathbf{S}_{ij}\|^2 \quad (3)$$

其中  $S$  是预定义的标签矩阵,  $f_1(\bullet)$  和  $f_2(\bullet)$  是两个不同的哈希函数。从技术上讲, 他们将样本  $x_i$  和  $x_j$  之间的相似度近似为和  $f_1(x_i)$  和  $f_2(x_j)$  之间的汉明距离。两个样本  $x_i$  和  $x_j$  越相似,  $f_1(x_i)$  和  $f_2(x_j)$  之间的汉明距离应该越小。与这些专注于单域检索的方法不同, 本文利用非对称哈希机制来解决 GTR 问题, 特别是其具有挑战性的 HeSDR 和 HeCDR 子任务。更重要的是, 由于在无监督场景中没有为  $S$  的构建提供标签信息, 因此我们设计了一种配备自适应二部图的新型距离相似度乘积函数, 以从数据中学习潜在的语义信息。

## 3 本文方法

### 3.1 图和二分图

假设  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$  是一个无向加权图, 具有顶点集  $\mathcal{V}(\mathcal{G})$  边集  $\mathcal{E}(\mathcal{G})$  和邻接矩阵  $\mathbf{W}$ , 其中实对称矩阵  $\mathbf{W}$  的每个元素测量一对顶点。一般来说, 亲和图的构建主要有三种方法: 1) 最近邻图 [7], 重点关注数据的局部邻域关系; 2) 语义图 [8], 利用标签的先验语义信息; 3)

自适应图 [9], 从数据中学习自适应图结构。另一方面, 二部图在不同的应用中得到了广泛的探索, 其定义如下。定义: 如何一个图  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$  满足  $\mathcal{V} = \mathcal{V}_s \cup \mathcal{V}_t$  且  $\mathcal{V}_s \cap \mathcal{V}_t = \emptyset$ , 那么这个图被称为二分图。注意一个二分图  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$  的亲和矩阵为  $W \in \mathbb{R}^{|\mathcal{V}_s| \times |\mathcal{V}_t|}$  不一定是对称的。

### 3.2 无监督 ATH

图 2 说明了所提出的 ATH 框架的核心思想, 该框架包含三个主要部分, 包括非对称哈希学习、跨域信息交互和域结构保存。在本节中, 首先考虑无监督的 ATH (ATH\_U), 其中源域和目标域都不包含标记样本。

1) 非对称哈希学习: 在 ATH 中, 利用两个不同的哈希函数  $f_t$  和  $f_s$  分别对目标样本和源样本进行编码, 如下所示:

$$\begin{aligned} \min_{\mathbf{B}_k, f_k} \mathcal{T}_1 &= \sum_{k \in \{s, t\}} \alpha_k \|\mathbf{B}_k - f_k(\mathbf{X}_k)\|_F^2 \\ \text{s.t. } \mathbf{B}_k &\in \{-1, 1\}^{r \times n_k}, \mathbf{B}_k \mathbf{1} = \mathbf{0} \end{aligned} \quad (4)$$

其中  $\alpha_k \in \{\alpha_s, \alpha_t\}$  是平衡参数,  $f_k \in \{f_s, f_t\}$  是域  $\mathcal{D}_k$  的哈希函数。我们施加一个约束  $\mathbf{B}_k \mathbf{1} = \mathbf{0}$ , 使每一位都有相同的机会成为 1 或 -1, 从而保持数据的最多信息。设计的优化目标 (4) 旨在对每个域  $\mathcal{D}_k$  的二进制代码  $B_k$  和哈希映射和哈希映射  $f_k(X_k)$  之间的拟合误差进行建模。需要注意的是,  $f_k(\bullet)$  可以是任何合适的映射函数, 例如线性函数和核函数。对于 ATH\_U, 我们只需利用一个简单而有效的线性函数:

$$f_k(x_{k,i}) = A_k^T x_{k,i} \quad (5)$$

其中  $A_k \in \{A_s, A_t\}$  是哈希投影。最值得注意的是,  $A_s \in \mathbb{R}^{d_s \times r}$  和  $A_t \in \mathbb{R}^{d_t \times r}$  可以具有相同的量级, 也可以不同。如果  $(d_t) = (d_s)$ , 则可用于 HoSDR 和 HoCDR。否则, 它会针对 HeSDR 和 HeCDR 进行定制。

2) 跨域信息交互: 我们考虑在目标域和源域之间进行直接的知识转移和融合, 无论这两个域是否共享相同的特征空间。显然, 像中那样最小化  $A_s$  和  $A_t$  之间的误差是不可行的, 因为跨域特征可能未对齐。根据定义 1, 我们观察到, 如果我们将源数据和目标数据视为一个完整的数据集, 那么我们可以自然地构造一个二分图  $\mathcal{G}_{st} = (\mathcal{V}_{st}, \mathcal{E}_{st}, \mathbf{W}_{st})$ , 其中  $\mathcal{V}_{st} = X_s \cup X_t$  和  $\mathbf{W}_{st} \in \mathbb{R}^{n_s \times n_t}$ 。然而, 如何确定边集  $\mathcal{E}_{st}$  和边权重  $W_{st}$  成为一个棘手的问题。请注意, 我们不能在 HeSDR 和 HeCDR 子任务中应用最近邻二部图, 因为无法直接计算跨域数据的异构特征之间的距离, 如图 2 所示。为了解决这个问题, 我们建议学习自适应二部图从数据中通过以下距离相似度乘积函数来表征跨域样本的亲合力:

$$\begin{aligned} \min_{\mathbf{W}_{st}, f_k} \nabla_2 &= \lambda \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} \mathbf{F}_{st,ij} \mathbf{W}_{st,ij} + \gamma \|\mathbf{W}_{st}\|_F^2 \\ \text{s.t. } \forall i, \quad \mathbf{W}_{st,i,:} &\geq 0, \quad \mathbf{W}_{st,i,:}^T \mathbf{1} = 1 \end{aligned} \quad (6)$$

$\lambda$  和  $\gamma$  是两个超参数,  $\mathbf{W}_{st,i}$  是  $\mathbf{W}_{st}$  的第  $i$  行,  $\mathbf{F}_{st,ij} = \|f_s(\mathbf{x}_{s,i}) - f_t(\mathbf{x}_{t,j})\|^2$  对距离进行建模 (或拟合跨域样本的哈希映射之间的错误)。  $\mathbf{W}_{st,ij}$  的值可以看作两个跨域样本  $x_{s,i}$  和  $x_{t,j}$  属于同一类的概率。因此, 我们设置约束  $\mathbf{W}_{st,i} \geq 0$ 。此外, 施加约束  $\mathbf{W}_{st,i} \mathbf{1} = 1$  以避免  $\mathbf{W}_{st,i,:}$  的所有元素均为零且正则化项  $\gamma \|\mathbf{W}_{st}\|_F^2$  提供均匀分布的先验概率。我们可以看到, 设计的距离相

似度乘积函数 (6) 使求解器能够同时学习跨域样本的哈希函数和亲和关系, 无论两个域的特征是否不一致。这个优化目标源于一个简单的直觉, 即两个跨域样本越相似 (即  $\mathbf{W}_{st,ij}$  越大), 相应的哈希映射差异应该越小 (即  $\mathbf{F}_{st,ij}$  越小)。非对称哈希函数和二部图的联合优化促进了之间的知识融合在跨域之间并使得在异构领域进行迁移训练成为可能。

3) 域结构保留在进行迁移训练时, 负迁移成为一个值得认真考虑的问题。为了缓解这个问题, 我们通过涉及域亲和图来保留单域数据的内在几何结构。具体来说, 对于每个域  $\mathcal{D}$ , 我们首先构造一个最近邻图  $\mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k, \mathbf{W}_k)$ , 其中  $\mathcal{V}_k = X_k$ ,  $\mathcal{E}_k$  定义如下:

$$\mathcal{E}_k = \{\langle \mathbf{x}_{k,i}, \mathbf{x}_{k,j} \rangle \mid i \in \mathcal{N}_{\eta_k}(j) \cup j \in \mathcal{N}_{\eta_k}(i)\} \quad (7)$$

其中  $\mathcal{N}_{\eta_k}(i)$  表示  $x_{k,i}$  和  $\eta_k$  个最近邻的集合。有了定义好的边集  $\mathcal{E}_k$ , 我们可以简单地设置边权重  $\mathbf{W}_k \in \mathbb{R}^{n_k \times n_k}$  如下:

$$\mathbf{W}_{k,ij} = \begin{cases} 1, & \text{if } \langle \mathbf{x}_{k,i}, \mathbf{x}_{k,j} \rangle \in \mathcal{E}_k \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

然后, 我们设计以下优化目标, 迫使每个域的相似样本在相应的映射空间中尽可能接近

$$\min_{f_k} \mathcal{T}_3 = \sum_{k \in \{s,t\}} \frac{\beta_k}{2} \sum_{i=1}^{n_k} \sum_{j=1}^{n_k} \mathbf{F}_{k,ij} \mathbf{W}_{k,ij} \quad (9)$$

$\beta_k \in \{\beta_s, \beta_t\}$  是平衡参数,  $F_{k,ij} = \|f_k(x_{k,i}) - f_k(x_{k,j})\|_2$  对单域样本的哈希映射之间的拟合误差进行建模。

#### 4) 总体目标函数

所提出的 ATH\_U 的总体目标函数公式如下:

$$\begin{aligned} \min_{\mathbf{B}_k, \mathbf{W}_{st} f_k} \quad & \mathcal{J} = \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3 \\ \text{s.t.} \quad & \mathbf{B}_k \in \{-1, 1\}^{r \times n_k}, \quad \mathbf{B}_k \mathbf{1} = \mathbf{0} \\ & \forall_i, \quad \mathbf{W}_{st,i,:} \geq 0, \quad \mathbf{W}_{st,i,:}^T \mathbf{1} = 1. \end{aligned} \quad (10)$$

(10) 可以通过精心设计的迭代算法进行优化, 如第 IV-A 节所示。

### 3.3 半监督 ATH

本节提供了 ATH 的半监督实现, 称为 ATH\_M, 其中源域中的所有样本都被很好地注释为 [3] 和 [10] 中的应用场景。ATH\_U 和 ATH\_M 目标函数的区别在于源图的构建。与 ATH\_U 不同, ATH\_M 对源域采用语义图。具体来说, ATH\_M 定义源图  $\mathcal{G}_s = (\mathcal{V}_s, \mathcal{E}_s, \mathbf{W}_s)$  的边集  $\mathcal{E}_s$  如下:

$$\mathcal{E}_s = \{\langle \mathbf{x}_{s,i}, \mathbf{x}_{s,j} \rangle \mid \mathbf{y}_{s,i} = \mathbf{y}_{s,j}\}. \quad (11)$$

ATH\_M 的边权重设置与 ATH\_U 类似。从 (11) 可以看出, ATH\_M 利用源域的语义信息进行传输训练, 从而提高了哈希编码器的性能。

### 3.4 有监督的 ATH 和核拓展

本节介绍了称为 ATH\_S 的 ATH 的监督实现, 它是针对源域和目标域中的所有样本都被很好标记的情况量身定制的。有了可用的标签信息, 我们不需要像 (6) 中那样优化

跨域知识转移的自适应二分图。相反，ATH\_S 主动构建语义二分图并设置其边集  $\mathcal{E}_{st}$  如下：

$$\mathcal{E}_{st} = \{ \langle \mathbf{x}_{s,i}, \mathbf{x}_{t,j} \rangle \mid \mathbf{y}_{s,i} = \mathbf{y}_{t,j} \}. \quad (12)$$

这样，跨域数据的语义相似性可以用于进行域对齐并提高模型的可迁移性。对于源图和目标图

$$\mathcal{G}_k = (X_k, \mathcal{E}_k, \mathbf{W}_k) \quad (13)$$

$\mathcal{D}_k$  构造边集  $\mathcal{E}_k$ ，如下所示：

$$\mathcal{E}_k = \{ \langle \mathbf{x}_{k,i}, \mathbf{x}_{k,j} \rangle \mid \mathbf{y}_{k,i} = \mathbf{y}_{k,j} \}. \quad (14)$$

如 (4) 中所述，我们可以使用核函数来实现  $f_k$ 。为了探索核函数的性能，我们进一步提出了 ATH\_S 的核扩展，称为 ATH\_K，其中哈希函数可以表示如下：

$$\hat{f}_k(\mathbf{x}_{k,i}) = \mathbf{A}_k^T \phi(\mathbf{x}_{k,i}) \quad (15)$$

其中  $\phi(\mathbf{x}_{k,i})$  是高斯核得到的  $m_k \in \{m_s, m_t\}$  维列向量

$$\phi(\mathbf{x}_{k,i}) = \left[ \exp\left(\frac{\|\mathbf{x}_{k,i} - \chi_{k,1}\|^2}{\sigma}\right), \dots, \exp\left(\frac{\|\mathbf{x}_{k,i} - \chi_{k,m_k}\|^2}{\sigma}\right) \right]^T \quad (16)$$

其中  $\{\chi_{k,j}\}_{j=1}^{m_k}$  是从域  $\mathcal{D}_k$  中随机选择的  $m_k$  个锚点,  $\sigma$  是核宽度

## 4 复现细节

### 4.1 与已有开源代码对比

本文的算法和实验没有参考其他源代码，均为自己复现。ATH\_S 算法的伪代码如 Algorithm 所示，输出的结果为 hash 投影矩阵，进而可以统一对齐特征分布：

---

**Algorithm 1** Optimization Algorithm for ATH\_S

---

**Input:** Training sets  $\mathbf{X}_k \in \{\mathbf{X}_s, \mathbf{X}_t\}$ , maximum iterations  $\text{Ite}$ , parameters  $\alpha_k, \beta_k, \lambda, r$ .

**Output:** Hashing projections  $\mathbf{A}_k \in \{\mathbf{A}_s, \mathbf{A}_t\}$ .

---

```

1 Construct affinity matrices  $\mathbf{W}_k \in \{\mathbf{W}_s, \mathbf{W}_t\}$ ;
2 Initialize  $\mathbf{A}_k$  by PCA projections;
3 Initialize  $\mathbf{B}_k$  by  $\mathbf{B}_k = \text{sign}(\mathbf{A}_k^T \mathbf{X}_k)$ ;
4 while converge or reach maximum iterations  $\text{Ite}$  do
5   | Update  $\mathbf{A}_k$  according to (17);
6   | Update  $\mathbf{M}_k$  by  $\mathbf{M}_k = \alpha_k \mathbf{A}_k^T \mathbf{X}_k$ ;
7   | for  $i = 1 : r$  do
8     | | Update  $\mathbf{B}_{k,ij}$  according to (18);
9   | end
10  | for  $i = 1 : n_s$  do
11    | | Update  $\mathbf{W}_{st,i}$  according to (19);
12  | end
13 end
```

---

$$\mathbf{A}_k = \Phi_k^{-1} \left( \alpha_k \mathbf{X}_k \mathbf{B}_k^T + \lambda \mathbf{X}_k \widetilde{\mathbf{W}}_{st} \widetilde{\mathbf{X}}_k^T \widetilde{\mathbf{A}}_k \right) \quad (17)$$

$$\mathbf{B}_{k,ij} = \begin{cases} 1, & \text{if } j \in \tau(i) \\ -1, & \text{otherwise} \end{cases} \quad (18)$$

$$\mathbf{V}_{st,i:} = \left( \frac{1 + \sum_{j=1}^{\eta} \frac{\lambda}{2\gamma_i} \widetilde{\mathbf{F}}_{st,ij}}{\eta} \mathbf{1} - \frac{\lambda}{2\gamma_i} \mathbf{F}_{st,i:} \right)_+ \quad (19)$$

## 4.2 数据集和比较算法

为了评估 ATH 框架的性能，在单域和跨域数据集上进行了实验。使用的数据集描述如下。

对于单域数据集来说使用 CIFAR-10 数据集，CIFAR-10 数据集包含 10 个类别的 60000 张图像，其中随机选择 59000 张图像（从每个类别采样 5900 张图像）进行训练，其余图像用于测试。

对于跨域数据集来说在 OfficeHome 上进行实验，这是一个著名的跨域基准测试，包含四个不同的域：1) 艺术图像 (A); 2) 剪贴画 (C); 3) 产品图像 (P); 4) 现实世界 (R)。为了简单起见，使用  $\# \rightarrow \&$  来表示源域为  $\#$ ，目标域为  $\&$ 。

使用的对比算法为 SDH 算法。

Dataset	Sample	Dimensions	Classes
OfficeHome	6866	2048	65
CIFAR-10	60000	512	10

表 1. 数据集描述

## 4.3 重建实验 ACC 与迭代次数的关系

显示了所提出的 ATH 算法在跨域 OfficeHome 上 A $\rightarrow$ P 任务上的准确性。正如图 3 所示，当算法迭代到第六次的。此外，随着迭代次数的增加，精度逐渐增加并最终达到稳定，这证明了所提出的 ATH 所学习到的映射空间的有效性。因此，ATH 算法可以在训练过程中实现正迁移，这是该方法在不同 GTR 子任务上优于其他比较方法的潜在原因。同时，图中也显示了在单域 CIFAR-10 数据集集中的表现优于跨域的性能，这也很符合预期的设想，因为跨域的学习难度是大于单域的学习的。



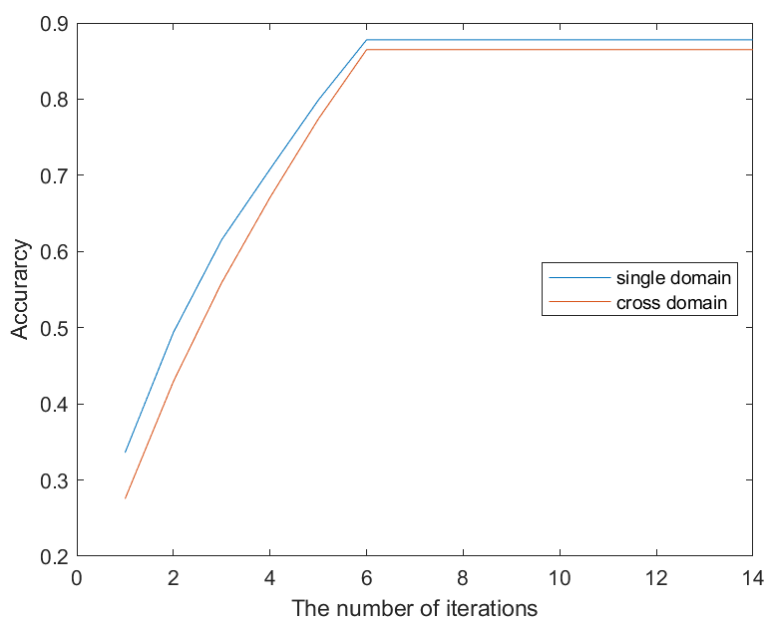


图 3. ACC 和迭代次数的关系

#### 4.4 重建实验模型的有效性

进行实验来证明所提出算法的有效性是否取决于训练样本的大小。具体来说，在 OfficeHome 跨域数据集和 CIFAR-10 单域数据集上进行了实验。在实验中，改变训练样本的数量并计算了测试集上的 MAP 结果。图 4 显示了 MAP 结果（16 位）随训练样本数量变化的曲线。很容易看出，ATH 算法的算法仅用大约 40% 的数据 (400) 就可以取得良好的性能。此外，当训练样本数量增加到一定程度时，ATH 算法的性能趋于稳定。

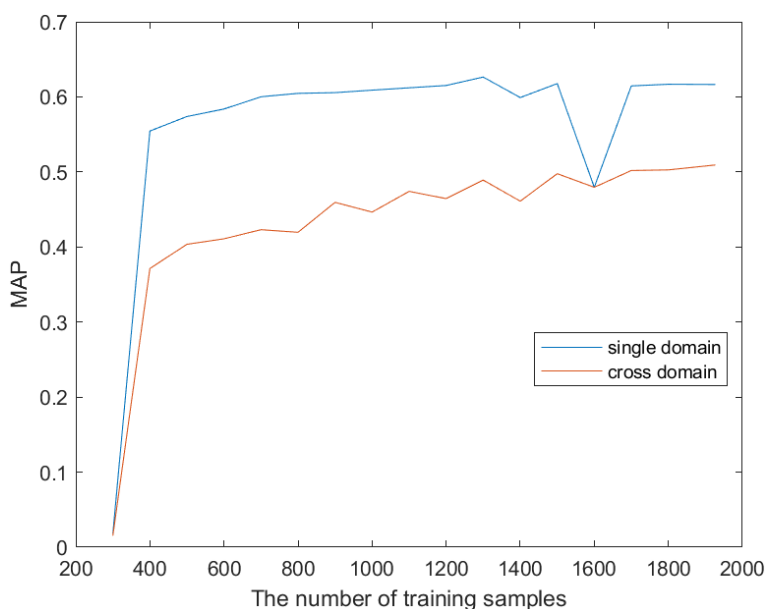


图 4. 样本数量与 MAP 的关系

#### 4.5 重建实验单域数据集的性能

表 1 提供了有监督哈希方法（包括 ATH\_S 算法和 SDK 算法）在单域基准数据集 CIFAR-10 数据集上的 MAP 结果（最高标记为红色）。正如表 2 表示，在单域设置中，基于相似性保留的方法 ATH\_S 总是 SDH 算法获得更好的性能，这强调保持成对样本之间相似关系的重要性。此外，虽然所提出的 ATH\_S 是为迁移检索量身定制的，但它在单域基准测试上也可以优于传统的哈希方法。潜在的原因是所提出的 ATH\_S 可以在训练阶段从数据中自适应地学习样本之间的相似关系。

Method	16	32	48	64	96	128
	CIFAR-10					
SDH	42.12	43.39	44.88	46.54	47.22	48.13
ATH_S	58.25	59.52	60.15	61.89	62.07	63.09

表 2. MAP 与编码长度关系

此外还提供了有监督哈希方法（包括 ATH\_S 算法和 SDK 算法）在单域基准数据集 CIFAR-10 的 PR 曲线（16 位）。正如图三所示，所提出的 ATH\_S 算法有着比 SDH 算法更佳的性能，这证明了 ATH\_S 框架对于 GTR 子任务的有效性和多功能性以及知识迁移学习的能力。

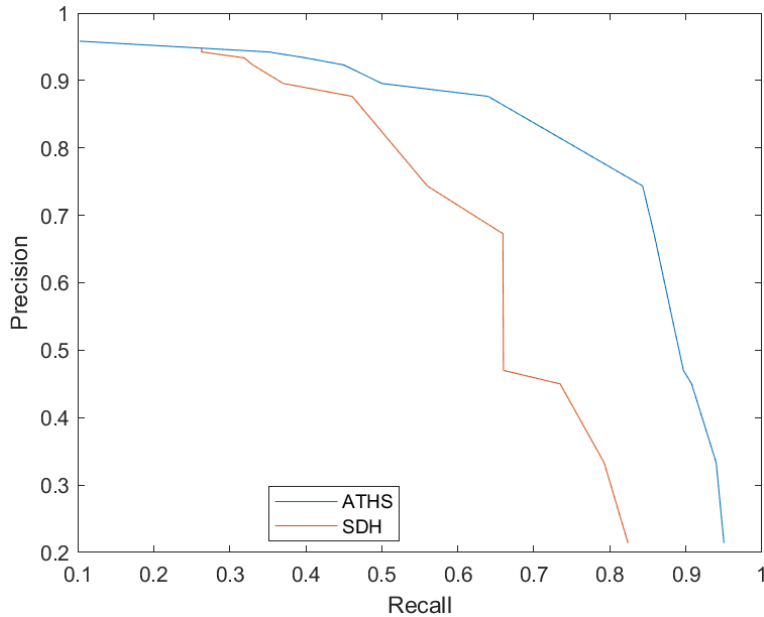


图 5. single-domain 的 RP 曲线

#### 4.6 重建实验跨域数据集的性能

表 3 提供了有监督哈希方法（包括 ATH\_S 算法和 SDK 算法）在跨域基准数据集 OfficeHome 数据集上的跨域同构和单域同构中的 MAP 结果（最高标记为红色）。正如表 3 表示，在跨域设置中，基于相似性保留的方法 ATH\_S 总是 SDH 算法获得更好的性能，这强调保持成对样本之间相似关系的重要性。原因是 ATH\_S 的方法可以在训练期间进行正迁

Method	HoSDR			HoCDR		
	P→A	R→C	P→C	C→A	A→P	C→P
SDH	60.45	48.31	50.12	52.44	40.58	60.26
ATH_S	76.40	70.95	64.64	66.12	90.32	79.15

表 3. MAP of cross-domain

移，不仅可以保证跨域之间的图结构关系可以自适应的学习，而且各个域内的结构信息也可以通过自适应学习得到精确的图结构关系。

此外还提供了有监督哈希方法（包括 ATH\_S 算法和 SDK 算法）在跨域基准数据集 OfficeHome 数据集的 PR 曲线（16 位）。正如图 6 所示，所提出的 ATH\_S 算法有着比 SDH 算法更佳的性能，这证明了 ATH\_S 框架对于 GITR 子任务的有效性和多功能性以及知识迁移学习的能力。

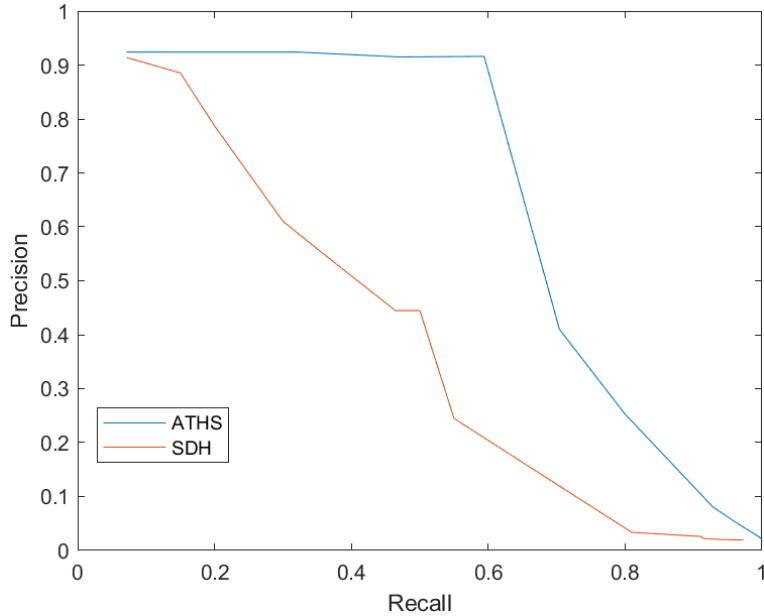


图 6. cross-domain 的 RP 曲线

#### 4.7 创新点

##### 1. 在 ATH 思想中引入锚点的思想

在 ATH 算法中跨域或者单域之间的亲和图构建的计算复杂度要高得多，因为该方法必须处理每对样本之间的相似性。此外，在预测过程中必须容忍  $n \times n$  ( $n$  是样本数) 相似度矩阵的逆，这对于大规模数据来说是无法承受的。因此，可以引入锚图的思想进入，锚点方法首先通过随机选择或聚类方法（例如 K-means）生成  $m$  个锚点，然后构建样本和锚点之间的亲和图。锚点的数量  $m$  明显小于原始样本的数量  $n$ ，并且利用原始数据和锚点构建亲和图只需要考虑  $O(mn)$  距离。

根据上述思想可得距离相似度内积目标函数为：

$$\begin{aligned}
\min_{\mathbf{W}_{st,f_k}} \nabla_2 &= \lambda \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} M_{st,ij} \mathbf{W}_{st,ij} + \gamma \|\mathbf{W}_{st}\|_F^2 \\
\text{s.t. } \forall_i, \quad &\mathbf{W}_{st,i:} \geq 0, \quad \mathbf{W}_{st,i:}^T \mathbf{1} = 1
\end{aligned} \tag{20}$$

其中,  $M_{st,ij}$  为不同域内的锚点之间所对应的哈希空间的欧氏距离。

## 2. 在 ATH 思想中引入 Possibilistic 思想

在 ATH 算法中, 对样本和周围邻居的相似性约束和进行了限制, 但是这将会产生一个问题, 那么就是如果样本  $x$  周围的另据有很多且都跟  $x$  距离很近, 但是由于和为 1 的限制, 那么将会导致样本  $x$  与周围的邻居的相似性值很小, 从而导致模型以为样本  $x$  是异常点或者是离群点。因此, 根据模糊理论的思想, 即默认每个数据点都在不同程度上属于每个类, 例如: 样本  $x$  属于类别 1 的隶属度为 0.9, 属于类别 2 的隶属度为 0.5, 属于类别 3 的隶属度为 0.4。类比上述思想可得样本  $x$  与邻居  $x_1$  的相似性为 0.8, 与样本  $x_2$  的相似性为 0.5。进而可以避免上述存在的问题。

根据上述思想去掉和为 1 的约束, 可得距离相似度内积目标函数为:

$$\begin{aligned}
\min_{\mathbf{W}_{st,f_k}} \nabla_2 &= \lambda \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} M_{st,ij} \mathbf{W}_{st,ij} + \gamma \|\mathbf{W}_{st}\|_F^2 \\
\text{s.t. } \forall_i, \quad &\mathbf{W}_{st,i:} \geq 0,
\end{aligned} \tag{21}$$

## 5 总结与展望

本文提出了一个更困难的图像迁移检索问题, 称为 GITER, 它对查询和检索集合的域和特征空间没有限制。为了解决 GITER 问题, 提出了一个基于非对称哈希策略和自适应二分图学习的 ATH 框架, 具有无监督/半监督/监督实现。对不同 GITER 子任务和标签设置进行的广泛实验证明了所提出的 ATH 框架的有效性。基于本文的思想, 提出了另外的二个想法。一是针对于处理大数据任务时会产生高度的时间和空间复杂性的问题引入锚图的思想加速模型的训练过程进而处理大数据任务。二是针对于原始模型是对样本之间的相似性进行单纯形约束而导致误导模型学习的问题引入 Possibilistic 思想放开和为 1 的约束, 将单纯形约束转化为仿射型约束, 进而可以来提高对离群点和噪音点的鲁棒性。后续将会对这两个模型进行实验以验证模型的正确性和可行性。

## 参考文献

- [1] Joey Tianyi Zhou, Heng Zhao, Xi Peng, Meng Fang, Zheng Qin, and Rick Siow Mong Goh. Transfer hashing: From shallow to deep. *IEEE Transactions on Neural Networks and Learning Systems*, 29(12):6191–6201, 2018.
- [2] Lei Zhang, Ji Liu, Yang Yang, Fuxiang Huang, Feiping Nie, and David Zhang. Optimal projection guided transfer hashing for image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(10):3788–3802, 2020.
- [3] Haifeng Xia, Taotao Jing, Chen Chen, and Zhengming Ding. Semi-supervised domain adaptive retrieval via discriminative hashing learning. In *Proceedings of the 29th ACM International Conference on Multimedia*, MM ’21, page 3853–3861, New York, NY, USA, 2021. Association for Computing Machinery.
- [4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.
- [5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.
- [6] Behnam Neyshabur, Payman Yadollahpour, Yury Makarychev, Ruslan Salakhutdinov, and Nathan Srebro. The power of asymmetry in binary hashing, 2013.
- [7] Zhihui Lai, Yudong Chen, Jian Wu, Wai Keung Wong, and Fumin Shen. Jointly sparse hashing for image retrieval. *IEEE Transactions on Image Processing*, 27(12):6147–6158, 2018.
- [8] Jianglin Lu, Zhihui Lai, Hailing Wang, Yudong Chen, Jie Zhou, and Linlin Shen. Generalized embedding regression: A framework for supervised feature extraction. *IEEE Transactions on Neural Networks and Learning Systems*, 33(1):185–199, 2022.
- [9] Jianglin Lu, Jingxu Lin, Zhihui Lai, Hailing Wang, and Jie Zhou. Target redirected regression with dynamic neighborhood structure. *Information Sciences*, 544:564–584, 2021.
- [10] Fuxiang Huang, Lei Zhang, Yang Yang, and Xichuan Zhou. Probability weighted compact feature for domain adaptive retrieval. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9579–9588, 2020.