属性一致的多模态实体对齐知识图谱表示学习

Qian Li 计算机科学与工程学院，BDBC，北京航空航天大学，北京，中国 liqian@act.buaa.edu.cn

Shu Guo 国家计算机网络应急技术处理协调中心，北京，中国 guoshu@cert.org.cn

Yangyifei Luo 计算机科学与工程学院，BDBC，北京航空航天大学，北京，中国 luoyangyifei@buaa.edu.cn

Cheng Ji 计算机科学与工程学院，BDBC，北京航空航天大学，北京，中国 jicheng@act.buaa.edu.cn

Lihong Wang 国家计算机网络应急技术处理协调中心，北京，中国 wlh@cert.org.cn

Jiawei Sheng 中国科学院信息工程研究所、中国科学院网络空间安全学院、中国科学院大学，中国北京 shengjiawei@iie.ac.cn

Jianxin Li*计算机科学与工程学院，北京航空航天大学，中关村实验室，北京，中国 lijx@act.buaa.edu.cn

arXiv:2304.01563v1 [cs.CL] 4 Apr 2023

摘要

多模态实体对齐（MMEA）旨在找出多模态知识图谱（MMKGs）之间的所有等价实体对。丰富的属性和邻近实体对于对齐任务具有重要价值，但现有工作忽视了上下文差异问题：在学习实体表示时，对齐实体在特定模态上具有不同数量的属性。本文提出了一种面向多模态实体对齐（MMEA）的新型属性一致性知识图谱表示学习框架（ACK-MMEA），通过融合一致性对齐知识来弥补上下文差异。首先，通过结合合并与生成算子进行多模态属性统一化，构建属性一致的知识图谱（ACKGs），使每个实体在每种模态中都具有且仅具有一个统一特征。随后，将ACKGs输入带有随机丢弃的关系感知图神经网络，以获得聚合关系表示和鲁棒实体表示。为了评估ACK-MMEA对实体对齐的促进作用，我们专门设计了一种同时用于实体与属性评估的联合对齐损失。在两个基准数据集上开展的大量实验表明，与竞争方法相比，我们的方法取得了优异的性能。

CCS 概念

• **计算方法** → **知识表示与推理**；**自然语言处理**；• **计算理论** →**语义与推理**。

关键词

实体对齐，多模态知识图谱表示

ACM 引用格式：

Qian Li、Shu Guo、Yangyifei Luo、Cheng Ji、Lihong Wang、Jiawei Sheng 和 Jianxin Li。2023年。《面向多模态实体对齐的属性一致性知识图谱表示学习》。载于*2023*年*ACM Web*会议论文集（*WWW '23*），*2023*年*5*月*1*日至*5*日，美国得克萨斯州奥斯汀。ACM，美国纽约州纽约市，11页。https://doi.org/10.1145/3543507.3583328

1 引言

知识图谱（KG）已成为一种广受欢迎的数据结构，用于以RDF三元组的形式表示事实知识。近年来，将多模态信息融入知识图谱的趋势日益显著，由此形成了多模态知识图谱（MMKG），可支持多种跨模态任务，例如推荐系统 [19, 28]和问答系统 [11, 27]。然而，多模态知识图谱往往存在覆盖率低和不完整的问题。为提高多模态知识图谱的覆盖率，一种可行的方法是采用多模态实体对齐（MMEA），通过整合文本和图像的属性信息，在不同的多模态知识图谱中识别等价实体对（即对齐种子）。这样，多模态知识图谱便能从其他知识图谱中获取有用的知识。

尽管多模态知识图谱中丰富的属性和邻近实体为多模态实体对齐 [14]提供了宝贵的证据，但多模态知识图谱不可避免的异质性使得从不同模态中学习和融合知识表示变得困难。一系列有效的

*通讯作者

允许为个人或课堂使用而免费制作本作品的全部或部分数字副本或纸质副本，但前提是不得出于营利或商业利益制作或分发副本，并且副本首页必须附有本声明及完整引文。本作品中由作者以外的其他方拥有的组件的版权必须得到尊重。经注明出处后，允许进行摘要。若要以其他方式复制、再版、发布到服务器或重新分发到列表，则需事先获得具体许可和/或支付费用。请通过 permissions@acm.org 申请许可。*WWW '23*，*2023*年*5*月*1*日至*5*日，美国得克萨斯州奥斯汀© 2023年版权所有者/作者(s)持有版权。出版权已授权给 ACM。ACM ISBN 978-1-4503-9416-1/23/04...$15.00 https://doi.org/10.1145/3543507.3583328

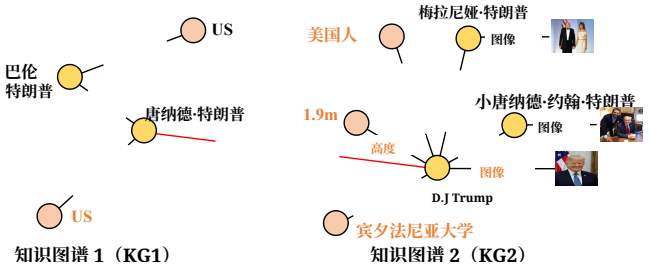


图1：知识图谱1与知识图谱2之间多模态实体对齐任务的示例。黄色和橙色圆圈分别表示实体节点和属性节点。

为克服这些挑战，研究者提出了多种方法，详细描述见附录A。PoE方法 [14] com- posed 通过拼接所有模态特征来构建实体表示，无法捕捉异构模态之间潜在的交互，因此限制了多模态实体对齐的性能。后续工作 [3, 7] 设计了多模态融合模块，以适当地整合属性和实体，从而根据聚合嵌入更好地预测对齐。所有这些方法都会利用实体关联的全部属性和邻近实体来学习实体表示。然而，它们忽略了实体对之间的上下文差异，进而限制了实体对齐的有效性。

上下文差距是指实体可能关联不同数量的属性，甚至缺少某些模态；由于信息冗余或缺失，这种情况不可避免。等价实体对之间的这种不一致会使对齐容易出错。图1展示了多模态实体对齐中的一个上下文差距示例。(a) 属性数量不同。知识图谱1中的实体唐纳德·特朗普 仅关联一个文本属性美国。由于上下文信息有限，难以确定实体唐纳德·特朗普的身份。相比之下，知识图谱2中的实体*D.J Trump* 包含更丰富的文本属性，因此其身份会更加具体；这导致现有基于聚合的方法会淡化文本属性（美国 和美国人之间）的相似度。文本属性数量不同造成的上下文差距，使得难以获得属性一致的对齐知识，也难以判断它们实际上指向同一个现实世界中的身份。(b) 缺少模态属性。缺失属性也会导致上下文差异问题，因为现有融合基线方法会忽略用于对齐的独特属性；例如，唐纳德·特朗普 缺少图像属性，导致*D.J Trump* 难以与其对齐。

为克服上述挑战，我们提出一种新颖的**多**模态**多**实体**对**齐**框**架，基于**属**性一致**知**识图谱表示学习，称为 ACK-MMEA[1]。具体而言，针对属性信息，我们设计了一种多模态属性统一化方法，为所有实体的每种模态获取具有统一属性的属性一致性知识图谱。也就是说，在属性一致的知识图谱中，每个实体在每种模态下都只有一个属性（即一个文本属性和一个图像属性）。为了生成这样的属性一致多模态知识图谱，我们设计了合并♀♀♀♀♀♀并分别为每个实体生成算子，以处理属性冗余和属性缺失问题。前者采用基于注意力的方法，将相关的多个属性压缩为一个，以过滤掉噪声属性；后者则在实体在特定模态下没有属性时扩展出一个新属性。我们进一步设计了 ConsistGNN，以强制对实体表示进行一致的属性聚合。给定上述属性一致的知识图谱中的一个关系三元组，ConsistGNN 首先通过同时整合每种模态（即实体/文本/图像）下的关系特征，获得聚合的关系表示。然后，使用关系感知实体编码器获得实体表示。由于新构建的属性一致的知识图谱可能会引入噪声，因此采用对邻居进行随机丢弃的方法，以生成更加鲁棒的表示。最后，我们为实体和属性评估设计了联合对齐损失。实验结果表明，我们的方法在两个多模态实体对齐基准数据集上均能取得优异的性能。我们的贡献总结如下：

- 我们提出了一种新颖的多模态实体对齐框架，通过利用属性一致性知识图谱表示学习方法来融入一致性对齐知识。据我们所知，这是首个解决多模态实体对齐任务中上下文差异问题的工作。
- 我们设计了一种多模态属性均匀化方法，通过合并与生成算子构建属性一致的多模态知识图谱，并提出 ConsistGNN 模型来聚合一致信息。
- 实验结果表明，该框架在两个公开多模态实体对齐数据集上取得了最先进的性能。

2 预备知识

我们首先给出多模态知识图谱（MMKG）和多模态实体对齐（MMEA）的定义如下。

多模态知识图谱的定义。 多模态知识图谱记作 $KG = (E, R, A)$，由实体之间的关系及其相关属性组成。具体而言，E，R，A 分别是实体、关系和多模态属性的集合，其大小分别为 n_E，n_R，n_A。我们假设 KG 有两类属性，即文本属性 T 和图像属性 I。

本文旨在解决多模态知识图谱实体对齐任务中的上下文差距问题。一般而言，实体对齐任务包括跨语言实体对齐和多源实体对齐 [30]。遵循 [3, 7, 13]，以往多模态实体对齐研究，本文聚焦于后者。

多模态实体对齐任务的定义. 多模态实体对齐任务 [3, 7, 13] 旨在识别两个多模态知识图谱中的一对实体是否等价。具体而言，给定两个多模态知识图谱 KG_1 和 KG_2，以及一对实体对齐种子 (v, v')，其中 v 和 v' 是 KG_1 和 KG_2 中的实体，多模态实体对齐旨在识别它们是否等价。其主要过程是学习两个多模态知识图谱中的实体表示，并计算对齐种子 (v, v') 之间的相似度。实体对齐种子的集合为 $S = \{(v, v') \mid v \in E, v' \in E', v \equiv v'\}$。

[1]源代码可在 https://github.com/xiaoqian19940510/ACK-MMEA 获取。

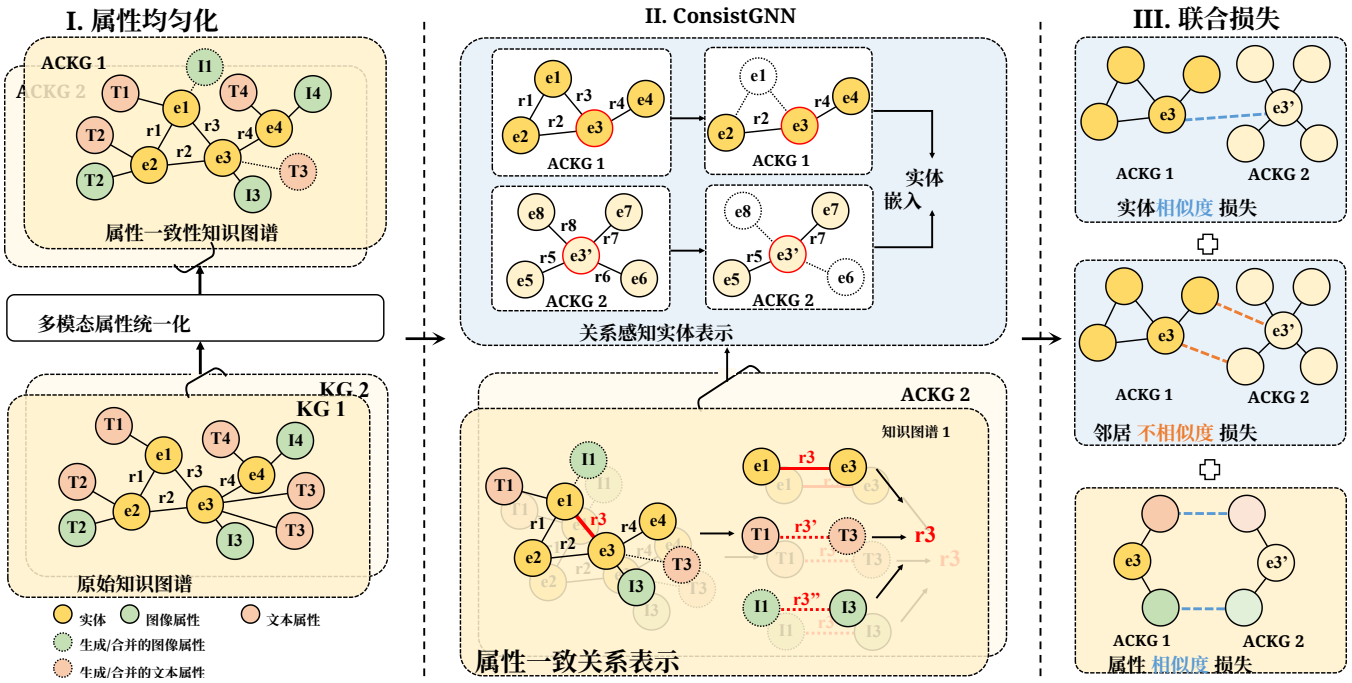


图 2：ACK-MMEA 的框架。(I) 对原始多模态知识图谱（KG_1 和 KG_2）进行多模态属性统一化，生成两个新的属性一致的多模态知识图谱（$ACKG_1$ 和 $ACKG_2$）。(II) ConsistGNN：带有 dropout 的关系感知图神经网络，用于聚合一致属性并学习鲁棒的实体表示。(III) 使用包含三个目标的联合对齐损失进行参数优化。

3 框架

本节介绍我们提出的框架 ACK-MMEA。如图 2 所示，ACK-MMEA 由以下三个模块组成：属性均匀化、ConsistGNN 和联合对齐损失。首先，属性均匀化模块为每个实体生成一致属性，其中每个实体在每种模态下分别具有一个属性。然后，ConsistGNN 将实体、属性和关系映射到共同表示空间，并通过聚合一致多模态属性，学习异构关系表示以及鲁棒的实体表示。最后，联合对齐损失结合实体相似度、属性相似度和邻居差异度的损失，对属性一致多模态知识图谱进行综合评估。

3.1 多模态属性统一化

为了更好地解决多模态知识图谱中属性不一致的问题，我们将原始多模态知识图谱 KG 划分为多个模态各自对应的知识图谱。

多模态知识图谱包含文本和图像属性，可划分为三个知识图谱 $\{KG_E,\ KG_T,\ KG_I\}$，分别对应实体图、文本图和图像图。对于每个划分后的知识图谱 KG_X，其中 $X \in \{E,\ T,\ I\}$，节点$\mathbf{E}_X$（实体、文本、图像）的表示为：

$$\mathbf{E}_X = \mathbf{F}_X \cdot \mathbf{W}_X, \tag{1}$$

其中$\mathbf{F}_X$是节点的初始表示。对于节点的每种模态，(a)在 KG_E 中，实体$\mathbf{F}_E \in \mathbb{R}^{n_E \times d_E}$ 由以下内容初始化。TransE 模型 [1], 的维度为 d_E；（b）在 KG_T 中，文本属性 $\mathbf{F}_T \in \mathbb{R}^{n_T \times d_T}$ 由 BERT [21], 初始化，其维度为 d_T；（c）在 KG_I 中，图像属性 $\mathbf{F}_I \in \mathbb{R}^{n_I \times d_I}$由 VGG16 [18], 初始化，其维度为 d_I。$\mathbf{W}_E \in \mathbb{R}^{d_E \times d}$、$\mathbf{W}_T \in \mathbb{R}^{d_T \times d}$、$\mathbf{W}_I \in \mathbb{R}^{d_I \times d}$是可学习的变换矩阵，将不同类型节点的初始表示映射到一个公共的 d 维空间。

如图3所示，原始多模态知识图谱中的实体属性并不一致。这种不一致是指，如第1节所述，每对实体在特定模态下的属性数量不同。这种上下文差异要求模型具备选择和生成有助于多模态实体对齐任务的属性的能力。然而，由于上下文差距没有特定的结构模式，且不同实体所面临的问题严重程度各异，因此很难确定实体对齐时应使用哪些属性以及使用多少属性。

为此，本文旨在寻求解决上下文差异问题的通用方案。我们在原始多模态知识图谱上实现以下两种统一化算子，将实体、属性及其关系映射到共同表示空间，同时确保它们之间的一致性。(a) 合并算子。我们通过基于注意力的机制，将每种模态的所有属性聚合为一个，以表示该模态下实体的统一特征，从而助力多模态实体对齐任务。(b) 生成算子。我们提出利用邻居属性生成缺失属性。合并算子与生成算子的组合缓解了上下文差异问题。

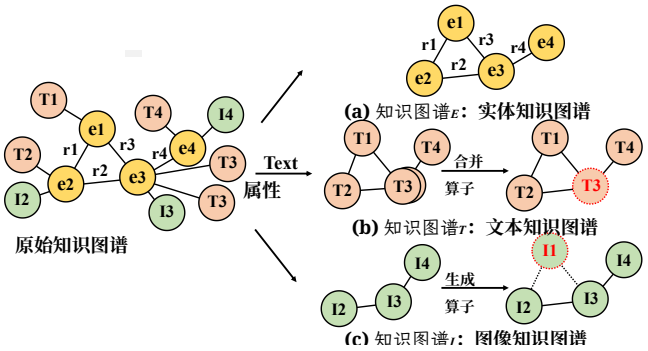


图3：多模态属性均匀化示意图。(a) 是仅包含实体的实体知识图谱。(b) 和 (c) 分别是文本知识图谱和图像知识图谱，它们依托 (a) 中的关系相连。我们使用合并与生成算子，使文本属性和图像属性实现均匀化。

合并算子。 对于多模态知识图谱，实体 $v \in \mathbf{E}$ 在 $\mathbf{E}_A$ 中可能有多个属性 $\{\mathbf{e}_{v,A,i}\}_{i=1}^{n_v}$，其中 $A \in \{T, I\}$， n_v 是 $\mathbf{h}_{v,A}$ 中属性的数量。为了聚合同一模态中各属性的信息，我们对每种模态执行合并算子。该算子通过可学习的图注意力 [15] 实现，能够以较低的权重舍弃冗余的上下文信息，使最终的一致属性有助于多模态实体对齐任务。因此，我们可以获得聚合属性 $\mathbf{E}_A^0$，该属性压缩了冗余的上下文信息：

$$\mathbf{E}_A^{(0)}[v] = \sigma(\sum_{i=1}^{n_v} \alpha_i \cdot \mathbf{W}_M \mathbf{e}_{v,A,i}), \tag{2}$$

其中， α_i 是第 i个属性的学习得到的注意力权重， $\mathbf{W}_M \in \mathbb{R}^{d \times d}$ 是可学习参数， σ 是 ReLU(·) 函数。

生成算子。对于多模态知识图谱，许多实体缺少特定模态的属性。观察发现，同一模态下邻居的属性通常能提供有用信息，用于生成目标实体的属性。直观来说，实体“唐纳德·特朗普”的图像属性与其子女的图像属性相似，这意味着它们的表示彼此接近。因此，为了弥补缺失属性造成的上下文差距，我们仅通过对一阶邻居属性进行平均聚合来生成属性$\mathbf{E}_{(0)\ A}$。

$$\mathbf{E}_A^{(0)}[v] = \sigma\left(\mathbf{W}_G \cdot \text{MEAN}\left(\{\mathbf{e}_{u,A} | u \in \mathcal{N}(v)\}\right), \tag{3}$$

其中 $\mathbf{W}_G \in \mathbb{R}^{d \times d}$ 是可学习的变换矩阵， $\mathbf{N}(v)$ 是实体 v 的一阶邻居集合， σ 是 ReLU(·) 函数，且 $\mathbf{e}_{u,A} = \mathbf{E}_A[u]$。

通过这种方式，可以缓解上下文差异，因为通过这两个统一化操作符，知识图谱中的实体属性是一致的。这样的属性一致的知识图谱（属性一致的知识图谱）有助于实体表示学习过程中的属性均衡，从而实现更准确的对齐。

3.2 ConsistGNN

我们进一步设计了一个名为 ConsistGNN 的图神经网络模型，分别基于属性一致性关系表示编码器和关系感知实体表示编码器推导关系和实体表示，从而在属性一致性知识图谱上实现一致的模态聚合。

在属性一致的知识图谱中，每个实体在每种模态下只有一个属性。因此，同一模态属性之间的关系与实体之间的关系有些相似，这是合理的。于是，我们将实体关系$\mathbf{R}_{(0)}$和属性关系$\mathbf{R}_{(0)\ A}$的表示分别定义如下：

$$\mathbf{R}^{(0)} = \mathbf{R} \cdot \mathbf{W}_0, \mathbf{R}_A^{(0)} = \mathbf{R}_A \cdot \mathbf{W}_{0,A}, \tag{4}$$

其中 $\mathbf{R}, \mathbf{R}_A \in \mathbb{R}^{n_r \times d_E}$ 是关系的初始表示，由 TransE 计算得出。具体而言，对于两个相连的实体（h, t），属性关系 $\mathbf{r}_{ht,A} = |\mathbf{e}_{t,A} - \mathbf{e}_{h,A}| \in \mathbf{R}_A$ 由尾部属性 $\mathbf{e}_{t,A}$ 和头部属性 $\mathbf{e}_{h,A}$计算得出。$\mathbf{W}_0, \mathbf{W}_{0,A} \in \mathbb{R}^{d_E \times d}$ 是可学习变换矩阵，用于将关系的初始表示映射到公共空间。

为了获得包含一致属性信息的实体表示，我们首先按如下方式初始化属性一致的知识图谱的实体表示：

$$\mathbf{E}_E^{(0)} = \sigma(\mathbf{E}_E \mathbf{W}_{c,E} + \sum_{A \in \{T,I\}} \mathbf{E}_A^{(0)} \mathbf{W}_{c,A}), \tag{5}$$

其中 $\mathbf{W}_{c,E}, \mathbf{W}_{c,A} \in \mathbb{R}^{d \times d}$ 是可学习变换矩阵。随后，我们将实体和关系的初始表示输入配备了属性一致关系表示编码器和关系感知实体表示编码器的 ConsistGNN。具体而言，我们按如下方式计算第 l层中节点 $\mathbf{E}_{(l)} = \{\mathbf{E}_{(l)\ E}, \mathbf{E}_{(l)\ A}\}$和关系 $\mathbf{R}_{(l)}$ 的表示：

$$\mathbf{E}^{(l)}, \mathbf{R}^{(l)} = \text{ConsistGNN}\left(\mathbf{E}^{(l-1)}, \mathbf{R}^{(l-1)}\right). \tag{6}$$

属性一致关系表示。 属性均匀化为每个实体的每种模态赋予唯一标识。两个实体之间的关系与它们的属性关系相近。因此，我们提出利用一致的属性信息，并借助属性关系进行实体关系学习。具体而言，我们提出一种属性一致关系表示编码器，以利用多模态属性信息，其中两个属性之间的关系由它们自身的组合表示：

$$\begin{aligned}\mathbf{R}^{(l)}[u, v] = \text{ReLU}(&\mathbf{W}_{E-E}^{(l)} \mathbf{r}_{uv}^{(l-1)} \\ &+ \sum_{A \in \{T,I\}} \mathbf{W}_{E-A}^{(l)} [\mathbf{e}_{u,A}^{(l-1)} || \mathbf{e}_{v,A}^{(l-1)}]),\end{aligned} \tag{7}$$

其中 $\mathbf{r}_{(l-1)uv} = \mathbf{R}_{(l-1)}[u, v]$， ReLU(·) 是激活函数。

关系感知实体表示. 为了使实体表示对生成的属性一致的知识图谱具有容错能力，我们对邻居进行随机丢弃 [29]，以提高实体表示的鲁棒性；该方法假设，部分实体缺失不会影响属性一致的知识图谱的语义含义。对于每个实体 v，我们随机丢弃一定比例的邻居实体，以及与这些实体相连的关系。邻居

每个实体 v 的集合因此更新为：

$$\begin{aligned}\mathcal{N}(v) &\leftarrow \text{drop}\left(\mathcal{N}(v);\rho\right)\\ &=\{u_i|u_i \in \mathcal{N}(v), p(i)=1, p(i)\sim \text{Ber}(1-\rho)\},\end{aligned} \tag{8}$$

其中，ρ 是随机丢弃率。drop(·) 是邻居随机丢弃函数，Ber(·) 是伯努利分布。最佳随机丢弃率为 0.35。随后，我们利用每个实体的关系和属性信息，通过邻居随机丢弃知识图谱更新实体表示。第 l层中的实体表示是

$$\mathbf{E}_E^{(l)}[v]=\mathbf{W}_{h,E}[\mathbf{e}_{v,E}^{(l-1)}||\frac{1}{D_v}\sum_{u\in\mathcal{N}(v)}[\mathbf{e}_{u,E}^{(l-1)}||\mathbf{r}_{uv}^{(l)}]], \tag{9}$$

其中，D_v 是实体 v的度。属性的更新是实体更新的特殊情况。第l层中的 $E_A^l[v]$ 是第(l-1)层中的属性表示$E_A^{l-1}[v]$ 和实体表示 $E_E^{l-1}[v]$ 的融合。

3.3 联合对齐损失

我们设计了一种联合对齐损失函数，利用实体相似度、属性相似度和邻居差异度的对齐损失，从多个视角评估属性一致多模态知识图谱。

对齐实体相似度.实体相似度约束损失如下：

$$\mathcal{L}_i^{EA}=\text{sim}(\mathbf{e}_i,\mathbf{e}_i')-\text{sim}(\mathbf{e}_i,\bar{\mathbf{e}}_i')-\text{sim}(\bar{\mathbf{e}}_i,\mathbf{e}_i'), \tag{10}$$

其中，$(\mathbf{e}_i,\mathbf{e}_{ri})$ 是 KG_1和 KG_2 中对齐种子 (v_i,v_{ri}) 的最终表示。$\mathbf{e}_i$和 $\mathbf{e}_{ri}$是这些种子的负样本。sim(·,·) 是余弦距离。

对齐属性相似度. 此外，为了使两个相邻实体的同类型属性之间的关系相似，我们设计了如下属性相似度约束损失：

$$\mathcal{L}_i^{attr}=\sum_{A\in\{T,I\}}\text{sim}(\mathbf{e}_{i,A},\mathbf{e}_{i,A}'). \tag{11}$$

对齐邻居差异度。 此外，为了更精确地对齐候选实体，KG_2 中的实体 v_{ri} 应与 KG_1中的实体 v_i 具有相似的特征，而 KG_2中 v_{ri}的邻居 N (v_{ri})应与 KG_1中的实体 v_i具有不同的特征。因此，受 [9, 25] 启发，我们引入邻居差异约束损失：

$$\mathcal{L}_i^{cont}=-\log\frac{\exp(\text{sim}(\mathbf{e}_i,\mathbf{e}_i')/\tau)}{\sum_{v_j'\in\mathcal{N}(v_i')}\exp(\text{sim}(\mathbf{e}_i,\mathbf{e}_j'))/\tau)}, \tag{12}$$

其中，v_rj是实体 v_{ri}的邻居，τ是温度系数。

总对齐损失 L 是三个损失函数的加权和：

$$\mathcal{L}=\lambda_1\mathcal{L}^{EA}+\lambda_2\mathcal{L}^{attr}+\lambda_3\mathcal{L}^{cont}, \tag{13}$$

其中 λ_1，λ_2，λ_3 是用于联合对齐损失的可学习超参数。

4 实验

4.1 数据集

我们在 FB15K-DB15K 和 FB15K-YAGO15K 数据集上进行了实验 [14], 这两个数据集是多模态实体对齐中最受欢迎的两个数据集。FB15K-DB15K 是 FB15K 和 DB15K 多模态知识图谱的实体对齐数据集[2]，包括 12,846 个对齐种子。FB15K-YAGO15K 是 FB15K 和 YAGO15K 知识图谱的实体对齐数据集，包括 11,199 个对齐种子。按照以往工作 [3, 7] 的建议，我们分别按照 2:8、5:5 和 8:2 的比例将这两个数据集划分为训练集和测试集。我们报告了在不同比例的对齐种子上进行评估时的官方平均倒数排名、Hits@1 和 Hits@10 指标。更多详情请参阅附录B。

4.2 对比方法

我们将本方法与六种实体对齐方法进行比较。这些方法最初聚合文本属性和关系信息。在这里，我们将由 VGG16 初始化的图像属性引入实体表示，并采用与文本属性相同的聚合方式：(1)**TransE** [1] 假设实体嵌入 v应接近属性嵌入 a与它们的关系 r之和。(2) **IPTransE** [32]是一种基于平移的方法，通过迭代式参数共享策略联合优化知识图谱中的实体和关系表示。(3) **GCN-align** [23]通过图卷积网络将各语言中的实体和属性映射到共同表示空间。(4) **SEA** [16] 通过对抗训练利用有标签和无标签实体，并结合图像属性。(5) **IMUSE** [8] 使用双变量回归来融合关系和多种属性。(6) **AttrGNN** [15]将知识图谱划分为多个子图，从而有效地建模各种类型的属性。

此外，我们将所提方法与四种多模态实体对齐方法进行比较。这些方法同样不引入名称属性，而是侧重于如何利用多模态属性：(7) **PoE** [14] 利用图像特征，并通过将实体语义与关系匹配来衡量可信度。(8) **PoE-rni** [14] 采用 PoE 中表现最佳的关系、数值字面量和图像属性。(9) **Chen 等人** [3] 设计了一个融合模块，以整合多模态属性。(10)**HEA** [7] 在双曲空间中刻画多模态知识图谱。

与这些方法不同，我们的方法会重新生成属性一致的多模态知识图谱，以统一属性信息，并通过采用联合损失的 ConsistGNN 学习更鲁棒的实体表示。

4.3 实现细节

对于所有基线方法，我们采用其文献中报告的最佳超参数。对于实体对齐基线方法（1–6），我们通过添加图像属性来复现其性能。对于多模态实体对齐基线方法（7–10），我们采用文献中报告的现有结果 [3, 7, 14]。

我们的模型基于 PyTorch（一种开源深度学习框架）实现。文本属性初始化使用 Hugging Face[3] 中的 BERT 版本 bert-base-uncased，图像属性初始化使用 VGG 版本 VGG164。GNN（GCN 和 GAT）层数为 2，训练轮次为 200，L2 正则化值

[2]https://github.com/mniepert/mmkb [3] https://github.com/huggingface/transformers [4] https://github.com/machrisaa/tensorflow-vgg

表 1：在不同实体对齐种子比例下，FB15K-DB15K（上）和 FB15K-YAGO15K（下）的主要实验结果。最佳结果以粗体标出，带下划线的数值为次优结果。"↑"表示相较于次优结果的提升，"-"表示该结果不可用。

方法	FB15K-DB15K (20%)			FB15K-DB15K（50%）			FB15K-DB15K（80%）		
	平均倒数排名（%）	Hits@1（%）	Hits@10（%）	平均倒数排名（%）	Hits@1（%）	Hits@10 (%)	平均倒数排名 (%)	Hits@1 (%)	Hits@10 (%)
TransE [1]	13.4	7.8	24.0	30.6	23.0	44.6	50.7	42.6	65.9
IPTransE [32]	9.4	6.5	21.5	28.3	21.0	42.1	46.9	40.3	62.7
GCN-align [23]	8.7	5.3	17.4	29.3	22.6	43.5	47.2	41.4	63.5
SEA [16]	25.5	17.0	42.5	47.0	37.3	65.7	50.5	51.2	78.4
IMUSE [8]	26.4	17.6	43.5	40.0	30.9	57.6	55.1	45.7	72.6
AttrGNN [15]	34.3	25.2	53.5	54.7	47.3	72.1	70.3	67.1	83.9
PoE [14]	17.0	12.6	25.1	53.3	46.4	65.8	72.1	66.6	82.0
PoE-rni [14]	28.3	23.2	39.0	44.2	38.0	55.7	55.8	50.2	64.1
Chen 等人 [3]	35.7	26.5	54.1	51.2	41.7	70.3	68.5	59.0	86.9
HEA [7]	-	12.7	36.9	-	26.2	58.1	-	41.7	78.6
ACK-MMEA（本文方法）	**38.7** (↑3.0)	**30.4** (↑3.9)	**54.9** (↑0.8)	**62.4** (↑7.7)	**56.0** (↑8.7)	**73.6** (↑1.5)	**75.2** (↑3.1)	**68.2** (↑1.1)	**87.4** (↑0.5)

方法	FB15K-YAGO15K（20%）			FB15K-YAGO15K（50%）			FB15K-YAGO15K（80%）		
	平均倒数排名(%)	Hits@1 (%)	Hits@10 (%)	平均倒数排名 (%)	Hits@1 (%)	Hits@10 (%)	平均倒数排名 (%)	Hits@1 (%)	Hits@10 (%)
TransE [1]	11.2	6.4	20.3	26.2	19.7	38.2	46.3	39.2	59.5
IPTransE [32]	8.4	4.7	16.9	24.8	20.1	36.9	45.8	40.1	60.2
GCN-align [23]	15.3	8.1	23.5	29.4	23.5	42.4	47.7	40.6	64.3
SEA [16]	21.8	14.1	37.1	38.8	29.4	57.7	60.5	51.4	77.3
IMUSE [8]	14.2	8.1	25.7	46.9	39.8	60.1	58.1	51.2	70.7
AttrGNN [15]	31.8	22.4	39.5	46.2	38.0	63.9	67.1	59.9	78.7
PoE [14]	15.4	11.3	22.9	41.4	34.7	53.6	63.5	57.3	74.6
PoE-rni [14]	33.4	25.0	49.5	49.8	41.1	66.9	57.2	49.2	70.5
Chen 等人 [3]	31.7	23.4	48.0	48.6	40.3	64.5	68.2	59.8	83.9
HEA [7]	-	10.5	31.3	-	26.5	58.1	-	43.3	80.1
ACK-MMEA（本文方法）	**36.0** (↑2.6)	**28.9** (↑3.9)	**49.6** (↑0.1)	**59.3** (↑9.5)	**53.5** (↑12.4)	**69.9** (↑3.0)	**74.4** (↑6.2)	**67.6** (↑7.7)	**86.4** (↑2.5)

为 0.0001，边距语法值为 1.0。对于超参数，最佳随机丢弃率 ρ 为 0.35，温度系数 τ 为 0.5，系数 $\lambda_1, \lambda_2, \lambda_3$ 分别为 5 、 3 和 2。对于学习率，我们采用步长为 0.001 的网格搜索方法。最优学习率为 0.001。全部超参数设置均通过 5 次网格搜索试验，在验证数据上进行调优。更多详情请参见附录 C。所有实验均在配备一块图形处理器（Tesla V100）的服务器上进行。我们方法的时间分析见附录 D。

4.4 主要结果

为验证 ACK-MMEA 的有效性，我们在表 1 中报告了整体平均结果。表中展示了在 FB15K-DB15K 和 FB15K-YAGO15K 数据集上，针对对齐种子训练集/测试集划分比例分别为 2:8、5:5 和 8:2 时的性能比较。

从表中我们可以观察到：1）在两个数据集上的三项指标方面，我们的属性一致性模型均优于所有实体对齐和多模态实体对齐方法的基线方法。具体而言，在所有训练数据比例下，就平均倒数排名而言，我们的模型在 FB15K-DB15K 上提升了 3.0% – 7.7%（平均提升 4%），在 FB15K-YAGO15K 上提升了 2.6% – 9.5%（平均提升 6%）。这表明我们的模型能够适应不同的训练资源比例，在多模态实体对齐任务中实现可靠的性能。2）与实体对齐基线方法（1-4）相比，尤其是在平均倒数排名和在 Hits@1 指标上，我们的模型在 FB15K-DB15K 和 FB15K-YAGO15K 上的平均提升分别达到 5% 和 9%，表现出更显著的提升。这表明，多模态一致属性统一化能够有效融入一致性对齐知识。3）与多模态实体对齐基线方法（5-8）相比，我们的模型基于新的属性一致多模态知识图谱设计了 ConsistGNN 模型；在平均倒数排名、Hits@1 和 Hits@10 方面，模型的平均提升分别为 5%、5% 和 1%。这是因为我们的方法融入了一致多模态属性和鲁棒的关系感知实体信息。4）对于两个数据集上的三种训练数据占比，在 Hits@1 指标上，我们的模型在 FB15K-DB15K 和 FB15K-YAGO15K 上的平均提升分别为 4.5% 和 8%。该指标表示仅预测标签与全局标签相等的占比。这表明，与基线方法相比，我们的方法更加准确，能够在只能预测一个结果时提供更多正确预测。上述观察均证明了 ACK-MMEA 框架的有效性。

4.5 模型变体讨论

为探究 ACK-MMEA 中各模块的有效性，我们进行了变体实验，结果见表2和图4。"↓"表示与 ACK-MMEA 相比的性能下降值。

表2：在 FB15K-DB15K (80%) 和 FB15K-YOGA15K (80%) 上的变体实验。“w/o”表示从完整模型中移除相应模块。“repl.”表示用另一个模块替换相应模块。“↓”表示与 ACK-MMEA 相比的性能下降值。

变体	FB15K-DB15K（80%） 平均倒数排名 (%)	Hits@1 (%)	Hits@10 (%)	△ 平均值 (%)
ACK-MMEA（本文方法）	**75.2**	**68.2**	**87.4**	-
不进行属性均匀化	73.2 (↓2.0)	63.9 (↓4.3)	82.8 (↓4.6)	↓3.6
不使用属性均匀化（合并算子）	74.0 (↓1.2)	65.1 (↓3.1)	83.5 (↓3.9)	↓2.7
不使用属性均匀化（生成算子）	74.2 (↓1.0)	65.6 (↓2.6)	84.1 (↓3.3)	↓2.3
不使用文本属性	73.4 (↓1.8)	66.2（↓2.0）	85.9 (↓1.5)	↓1.7
不含图像属性	72.5 (↓2.7)	65.8 (↓2.4)	86.2 (↓1.2)	↓2.1
替代。图卷积网络	73.5 (↓1.7)	66.6（↓1.6）	82.7 (↓4.7)	↓2.6
repl. 图注意力网络	74.1 (↓1.1)	65.5 (↓2.7)	83.1 (↓4.3)	↓2.7
不使用随机丢弃	72.7 (↓2.5)	66.0 (↓2.2)	84.6 (↓2.8)	↓2.5
替换。随机替换	71.2 (↓4.0)	64.9 (↓3.3)	83.8 (↓3.6)	↓3.6
无属性相似性损失	74.0 (↓1.2)	68.1（↓0.1）	85.7（↓1.7）	↓1.0
不使用邻居不相似性损失	73.6（↓1.6）	67.7（↓0.5）	86.9（↓0.5）	↓0.8

变体	FB15K-YOGA15K（80%） 平均倒数排名（%）	Hits@1（%）	Hits@10（%）	△ 平均（%）
ACK-MMEA（本文方法）	**74.4**	**67.6**	**86.4**	-
不使用属性均匀化	73.6 (↓0.8)	64.2 (↓3.4)	84.3 (↓2.1)	↓2.1
不使用属性均匀化（合并算子）	73.5 (↓0.9)	65.3 (↓2.3)	83.9 (↓2.5)	↓1.9
不使用属性均匀化（生成算子）	74.1 (↓0.3)	66.0 (↓1.6)	84.5 (↓1.9)	↓1.2
无文本属性	73.9（↓0.5）	65.8 (↓1.8)	84.7 (↓1.7)	↓1.3
不含图像属性	73.6（↓0.8）	65.7（↓1.9）	84.6（↓1.8）	↓1.5
替代图卷积网络	73.0 (↓1.4)	66.3 (↓1.3)	84.3（↓2.1）	↓1.6
替换图注意力网络	73.8（↓0.6）	65.9（↓1.7）	83.9（↓2.5）	↓1.6
不使用随机丢弃	72.3 (↓2.1)	65.7 (↓1.9)	84.1 (↓2.3)	↓2.1
替换。随机替换	70.9 (↓3.5)	64.2 (↓3.4)	83.1 (↓3.3)	↓3.4
不使用属性相似性损失	72.8 (↓1.6)	66.7 (↓0.9)	85.3 (↓1.1)	↓1.2
不使用邻居不相似性损失	73.2 (↓1.2)	67.0 (↓0.6)	86.1 (↓0.3)	↓0.7

从表2中可以观察到：1）属性均匀化对使用原始属性的影响往往更为显著。我们认为，这是因为一致属性为实体对齐提供了更多线索。2）将 ConsistGNN 替换为 GCN、GAT，或不对邻居进行随机丢弃或随机替换，都会导致性能显著下降。这表明 ConsistGNN 能够捕获更有效的一致属性和关系信息。3）属性相似度和邻居不相似性损失的影响往往较为显著。这是因为一致属性能够弥合上下文差异，而邻居损失则引导我们的模型学习鲁棒表示。4）当我们移除所有图像属性（“不使用图像属性”）时，我们的方法在 FB15K-DB15K 和 FB15K-YAGO15K 上的平均性能分别下降 2.1% 和 1.5%。当移除属性均匀化模块（“不使用属性均匀化”）时，平均性能分别下降 3.6% 和 2.1%。这表明图像属性能够提升模型性能，并且我们的方法通过捕获更多对齐知识有效地利用了图像属性。以上观察结果均证明了模型中各个组件的有效性。

为了进一步研究多模态属性对所有对比方法的影响，我们通过删除不同模态的属性来报告结果，如图4所示。从图中可以观察到：1）不含文本或图像属性的变体在所有评估指标上都显著下降，这表明多模态属性对于实体对齐任务是必要且有效的。2）删除所有多模态属性对我们模型的影响较小。我们认为，这是因为对邻居进行随机丢弃以及邻居不相似性损失有助于获得更好的实体表示。3）与其他基线方法相比，无论使用全部多模态属性还是舍弃其中一部分，我们的模型都取得了更好的结果。这表明我们的模型充分利用了现有的多模态属性，并且一致属性对于多模态实体对齐任务是有效的。4）当我们删除所有属性（标记为“删除所有属性”）时，我们的方法在平均倒数排名方面下降了1%–5%，但与所有基线方法相比仍表现最佳。这表明，通过关系感知实体表示，我们的模型使实体表示具备容错能力，并通过差异约束损失实现了对候选实体更精准的对齐。5）当我们

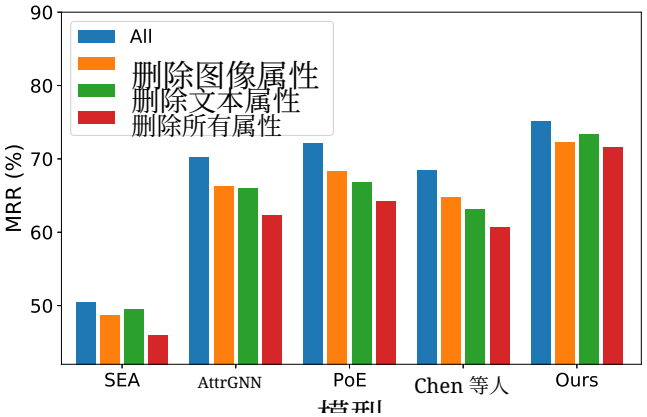


图4：在 FB15K-DB15K（80%）上删除属性的结果。“Del.”表示删除相应的属性。

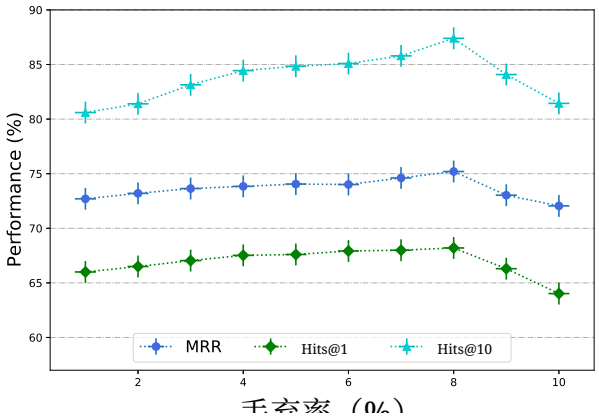


图5：对丢弃率 ρ 的分析。

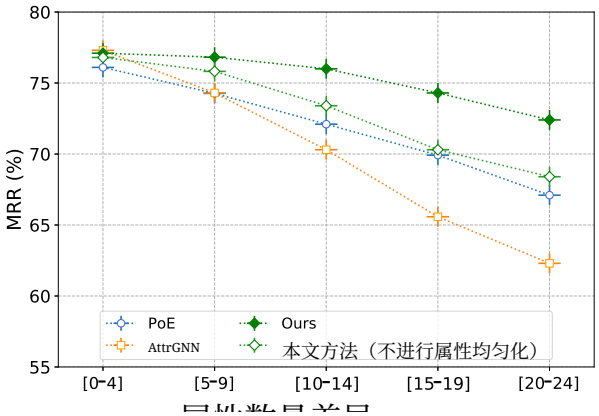


图6：属性数量差异的影响。

删除全部图像属性（“删除图像属性”），即将原始多模态知识图谱转换为知识图谱后，我们的方法优于其他基线方法。这表明，我们的模型通过属性一致关系表示编码器和关系感知实体表示编码器，融入了一致性对齐知识。所有这些观察结果都证明了所构建的属性一致多模态知识图谱和 ConsistGNN 的有效性。

4.6 丢弃率的影响

我们研究了随机丢弃率对邻居的影响。图5展示了在 FB15K-DB15K（80%）上，采用不同超参数设置的 ρ 的指标值。随着丢弃率增加，平均倒数排名、Hits@1 和 Hits@10 逐渐上升，并在达到最优值后开始下降。当邻居的丢弃率达到 35% 时，模型性能达到峰值，这反映了学习鲁棒实体表示的有效性。这表明随机丢弃能够提升容错能力，并促进一致的属性聚合。

4.7 属性数量的影响

我们研究了对齐种子之间不同程度的上下文差距所带来的影响。为此，我们选择具有相同数量图像属性的实体对齐种子，并改变 $[0, 24]$ 中实体对的文本属性数量差距。图6展示了 FB15K-DB15K (80%) 上属性差距变化情况下不同模型的性能。从图中可以观察到：1）随着对齐种子之间差距的增大，所有方法的性能逐渐下降。主要原因是，实体种子之间的差距越大，实体匹配就越困难。这再次验证了我们在引言中提出的直觉。2）与基线方法相比，我们模型的性能下降较为缓慢，证明了我们的方法在处理上下文差距问题方面的优越性。所有观察结果表明，我们的方法能够减小上下文差距的影响。

5 结论

本文提出了一种新颖的多模态实体对齐框架，即 ACK-MMEA。通过多模态属性统一化，该框架生成属性一致多模态知识图谱，其中每个实体在每种模态下仅包含一个属性。我们进一步提出 ConsistGNN，以整合一致的多模态属性，并获得聚合关系表示和鲁棒实体表示。为评估属性一致多模态知识图谱，我们设计了包含三个目标的联合对齐损失。我们的工作克服了实体对之间因信息冗余和属性缺失造成的上下文差异。实验结果表明，我们的方法能够解决上下文差异问题。然而，属性生成算子会引入噪声数据。未来，我们将研究如何避免噪声数据对多模态实体对齐任务的影响。

致谢

我们感谢匿名审稿人提出的富有洞察力的评论和建议。Jianxin Li 是通讯作者。本文作者获得了国家自然科学基金委员会项目（项目编号：U20B2053、62106059）以及北京航空航天大学博士生学术卓越基金的资助。

参考文献

[1] Antoine Bordes、Nicolas Usunier、Alberto García-Durán、Jason Weston 和 Oksana Yakhnenko。2013年。用于多关系数据建模的嵌入翻译。载于神经信息处理系统进展第*26*卷：*2013*年神经信息处理系统第*27*届年度会议。会议论文集，会议于*2013*年*12*月*5*日至*8*日在美国内华达州太浩湖举行。2787–2795。https://proceedings.neurips.cc/paper/2013/hash/1cecc7a77928ca8133fa24680a88d2f9-Abstract.html[2] Yixin Cao、Zhiyuan Liu、Chengjiang Li、Zhiyuan Liu、Juanzi Li 和 Tat-Seng Chua。2019年。用于实体对齐的多通道图神经网络。载于计算语言学协会第*57*届会议论文集，*ACL 2019*，意大利佛罗伦萨，*2019*年*7*月*28*日至*8*月*2*日，第*1*卷：长篇论文。计算语言学协会，1452 1461 10 18653 1 19 1140 – 。[3] https://doi.org/ 。/v /p - Liyi Chen、Zhi Li、Yijun Wang、Tong Xu、Zhefeng Wang 和 Enhong Chen。2020年。MMEA：多模态知识图谱的实体对齐。载于知识科学、工程与管理——第*13*届国际会议，*KSEM 2020*，中国杭州，*2020*年*8*月*28*日至*30*日，论文集，第一部分（计算机科学讲义，第*12274*卷）。施普林格，134–147。https://doi.org/10.1007/978-3-030-55130-8 12_[4] Liyi Chen、Zhi Li、Tong Xu、Han Wu、Zhefeng Wang、Nicholas Jing Yuan 和 Enhong Chen。2022年。用于实体对齐的多模态孪生网络。载于*KDD '22*：第*28*届*ACM SIGKDD*知识发现与数据挖掘会议，美国华盛顿哥伦比亚特区，*2022*年*8*月*14*日至*18*日。ACM，118–126。https: //doi.org/ 10 1145 3534678 3539244 。/ 。Muhao Chen 、 Yingtao Tian 、 Kai-Wei Chang 、[5] Steven Skiena 和 Carlo Zaniolo。2018年。协同训练知识图谱与实体描述的嵌入以实现跨语言实体对齐。载于第二十七届人工智能国际联合会议论文集，*IJCAI 2018*，*2018*年*7*月*13*日至*19*日，瑞典斯德哥尔摩。ijcai.org，3998 4004 10 24963 2018 556 – 。[6] https://doi.org/ 。/ijcai. / Muhao Chen、Yingtao Tian、Mohan Yang 和 Carlo Zaniolo。2017。。用于跨语言知识对齐的多语言知识图谱嵌入。载于第二十六届人工智能国际联合会议论文集，*IJCAI 2017*，澳大利亚墨尔本，*2017*年*8*月*19*日至*25*日。ijcai.org，1511–1517。https://doi.org/ 10 24963 2017 209 。/ijcai. / Hao Guo、Jiuyang Tang、Weixin Zeng、Xiang Zhao 和 Li Liu。2021年。双曲空间中的多模态实体对齐。神经计算 461 (2021)，598–607。https://doi.org/10.1016/j.neucom.2021.03.132[8] Fuzhen He、Zhixu Li、Qiang Yang、An Liu、Guanfeng Liu、Pengpeng Zhao、Lei Zhao、Min Zhang 和 Zhigang Chen。2019年。利用属性三元组和关系三元组进行无监督实体对齐。载于高级应用数据库系统——第*24*届国际会议，*DASFAA 2019*，泰国清迈，*2019*年*4*月*22*日至*25*日，论文集，第一部分（计算机科学讲义，第*11446*卷）。施普林格，367 382 10 1007 978 3 030 18576 3_22 – 。https://doi.org/ 。/ - - - - [9] Cheng Ji 、 Jianxin Li 、 Hao Peng 、 Jia Wu 、 Xingcheng Fu 、 Qingyun Sun 和 Philip S. Yu。2023年。无偏且高效的自监督增量对比学习。*CoRR* abs/2301.12104 (2023)。https://doi.org/10.48550/arXiv.2301.12104 arXiv:2301.12104[10] Jin Jiang、Mohan Li 和 Zhaoquan Gu。2021年。知识图谱中基于嵌入翻译的实体对齐综述。载于第六届*IEEE*网络空间数据科学国际会议，*DSC 2021*，中国深圳，*2021*年*10*月*9*日至*11*日。IEEE，187–194。https://doi.org/10.1109/DSC53577.2021.00033[11] Ke Liang、Lingyuan Meng、Meng Liu、Yue Liu、Wenxuan Tu、Siwei Wang、Sihang Zhou、Xinwang Liu 和 Fuchun Sun。2022年。不同类型知识图谱上的推理：静态、时序与多模态。*CoRR* abs/2212.05767 (2022)。https://doi.org/10.48550/arXiv.2212.05767 arXiv:2212.05767[12] Zhenxi Lin、Ziheng Zhang、Meng Wang、Yinghui Shi、Xian Wu 和 Yefeng Zheng。2022年。用于实体对齐的多模态对比表示学习。*CoRR* abs/2209.00891 (2022)。https://doi.org/10.48550/arXiv.2209.00891 arXiv:2209.00891[13] Fangyu Liu、Muhao Chen、Dan Roth 和 Nigel Collier。2021年。用于（无监督）实体对齐的视觉枢轴。载于第三十五届*AAAI*人工智能会议，*AAAI 2021*；第三十三届人工智能创新应用会议，*IAAI 2021*；第十一届人工智能教育进展研讨会，*EAAI 2021*，线上会议，*2021*年*2*月*2*日至*9*日。AAAI出版社，4257 4266 16550 – 。[14] https://ojs.aaai.org/index.php/AAAI/article/view/ Ye Liu、Hui Li、Alberto García-Durán、Mathias Niepert、Daniel Oñoro-Rubio 和 David S. Rosenblum。2019年。MMKG：多模态知识图谱。载于语义网——第*16*届国际会议，*ESWC 2019*，斯洛文尼亚波尔托罗日，*2019*年*6*月*2*日至*6*日，论文集（计算机科学讲义，第*11503*卷）。施普林格，459 474 10 1007 978 3 030 21348 0_30 – 。https://doi.org/ 。/ - - - - [15] Zhiyuan Liu、Yixin Cao、Liangming Pan、Juanzi Li 和 Tat-Seng Chua。2020年。探索并评估实体对齐的属性、值和结构。载于*2020*年自然语言处理实证方法会议论文集，*EMNLP 2020*，线上，*2020*年*11*月*16*日至*20*日。计算语言学协会，6355 6364 10 18653 1 2020 515 – 。[16] https://doi.org/ 。/v / .emnlp-main. Shichao Pei、Lu Yu、Robert Hoehndorf 和 Xiangliang Zhang。 。通过具备感知能力的知识图谱嵌入实现半监督实体对齐。2019。度差。在万维网大会，*WWW 2019*，美国加利福尼亚州旧金山，*2019*年*5*月*13*日至*17*日。ACM，3130–3136。https://doi.org/10.1145/ 3308558.3313646[17] Xiaofei Shi 和 Yanghua Xiao。2019。用于精确跨语言实体对齐的多重映射关系建模。载于*2019*年自然语言处理实证方法会议与第九届自然语言处理国际联合会议论文集，*EMNLP-IJCNLP 2019*，中国香港，*2019*年*11*月*3*日至7日。计算语言学协会，813–822。https://doi.org/ 10 18653 1 19 1075 。/v /D - Karen Simonyan 和 Andrew Zisserman[18]。2015。用于大规模图像识别的超深卷积网络。载于第三届国际学习表征会议，*ICLR 2015*，*2015*年*5*月7日至*9*日，美国加利福尼亚州圣迭戈，会议论文集。http://arxiv.org/abs/1409.1556[19] Rui Sun、Xuezhi Cao、Yan Zhao、Junchen Wan、Kun Zhou、Fuzheng Zhang、Zhongyuan Wang 和 Kai Zheng。2020。用于推荐系统的多模态知识图谱。载于*CIKM '20*：第*29*届*ACM*信息与知识管理国际会议，线上会议，爱尔兰，*2020*年*10*月*19*日至*23*日1405 1414 10 1145 3340531 3411947-。ACM，–。https://doi.org/。/。[20] Zequn Sun、Chengming Wang、Wei Hu、Muhao Chen、Jian Dai、Wei Zhang 和 Yuzhong Qu。2020。采用门控多跳邻域聚合的知识图谱对齐网络。载于第三十四届*AAAI*人工智能会议，*AAAI 2020*，第三十二届人工智能创新应用会议，*IAAI 2020*，第十届*AAAI*人工智能教育进展研讨会，*EAAI 2020*，*2020*年*2*月*12*日222 229 5354美国纽约州纽约市，-。–。https://aaai.org/ojs/index.php/AAAI/article/view/[21] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N. Gomez、Lukasz Kaiser 和 Illia Polosukhin。2017。注意力就是一切。载于神经信息处理系统进展第*30*卷：*2017*年*12*月*4*日至*9*日举行的神经信息处理系统年会，美国加利福尼亚州长滩。5998–6008。https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html[22] Zeshi Wang、Mohan Li 和 Zhaoquan Gu。2021。基于图卷积神经网络的实体对齐综述。载于第六届*IEEE*网络空间数据科学国际会议，*DSC 2021*，*2021*年*9*月*11*日144 151 10 1109 53577 2021中国深圳，*10*月-。IEEE，–。https://doi.org/。/DSC。。00027 2018 Zhichun Wang、[23] Qingsong Lv、Xiaohan Lan 和 Yu Zhang。。基于图卷积网络的跨语言知识图谱对齐。载于*2018*年自然语言处理实证方法会议论文集，比利时布鲁塞尔，*2018*年*10*月*31*日至*11*月*4*日。349–357。https://doi.org/10. 18653/v1/d18-1032[24] Zhichun Wang、Qingsong Lv、Xiaohan Lan 和 Yu Zhang。2018。基于图卷积网络的跨语言知识图谱对齐。载于*2018*年自然语言处理实证方法会议论文集。349–357。[25] Jiancan Wu、Xiang Wang、Fuli Feng、Xiangnan He、Liang Chen、Jianxun Lian 和 Xing Xie。2021。用于推荐的自监督图学习。载于*SIGIR '21*：第*44*届*ACM*信息检索研究与发展国际会议，线上会议，加拿大，*2021*年*7*月*11*日至*15*日。ACM，726–735。https://doi.org/ 10 1145 3404835 3462862 。/。Yuting Wu、Xiao Liu、Yansong Feng、[26] Zheng Wang 和 Dongyan Zhao。2020。用于实体对齐的邻域匹配网络。载于计算语言学协会第*58*届年会论文集，*ACL 2020*，线上会议，*2020*年*7*月*5*日至*10*日。计算语言学协会，6477–6487。https://doi.org/10.18653/v1/2020.acl-main.578[27] Fei Xia、Bin Li、Yixuan Weng、Shizhu He、Kang Liu、Bin Sun、Shutao Li 和 Jun Zhao。2022。LingYi：基于多模态知识图谱的医疗对话式问答系统。*CoRR* abs/2204.09220（2022）。https: //doi.org/ 10 48550 2204 09220 2204 09220 。/arXiv.。arXiv:。Guohai Xu、[28] Hehong Chen、Feng-Lin Li、Fu Sun、Yunzhou Shi、Zhixiong Zeng、Wei Zhou、Zhongzhou Zhao 和 Ji Zhang。2021。AliMe MKG：用于直播电商的多模态知识图谱。载于*CIKM '21*：第*30*届*ACM*信息与知识管理国际会议，线上会议，澳大利亚昆士兰，*2021*年*11*月*1*日至*5*日。ACM，4808–4812。https: //doi.org/10.1145/3459637.3481983[29] Yuning You、Tianlong Chen、Yongduo Sui、Ting Chen、Zhangyang Wang 和 Yang Shen。2020。采用增强的图对比学习。载于神经信息处理系统进展第*33*卷：神经信息处理系统年会，*NeurIPS 2020*，*2020*年*12*月*6*日至*12*日，线上。https://proceedings.neurips.cc/paper/2020/hash/3fe230348e9a12c13120749e3f9fa4cd-Abstract.html[30] Kaisheng Zeng、Chengjiang Li、Lei Hou、Juanzi Li 和 Ling Feng。2021。知识图谱实体对齐综合综述。*AI Open* 2（2021），1–13。https://doi.org/10.1016/j.aiopen.2021.02.002[31] Qingheng Zhang、Zequn Sun、Wei Hu、Muhao Chen、Lingbing Guo 和 Yuzhong Qu。2019。用于实体对齐的多视图知识图谱嵌入。载于第二十八届国际人工智能联合会议论文集，*IJCAI 2019*，中国澳门，*2019*年*8*月*10*日至*16*日。ijcai.org，5429–5435。https://doi.org/10.24963/ijcai.2019/754

[32] Hao Zhu、Ruobing Xie、Zhiyuan Liu 和 Maosong Sun。2017 年。通过联合知识嵌入进行迭代实体对齐。载于 第二十六届人工智能国际联合会议论文集，*IJCAI 2017*，澳大利亚墨尔本，*2017*年*8*月*19*日至*25*日。ijcai.org，4258–4264。https://doi.org/10.24963/ijcai. 2017/595[33] Xiangru Zhu、Zhixu Li、Xiaodan Wang、Xueyao Jiang、Penglei Sun、Xuwu Wang、Yanghua Xiao 和 Nicholas Jing Yuan。2022年。多模态知识图谱构建与应用：综述。*CoRR* abs/2202.05786（2022年）。arXiv:2202.05786 https://arxiv.org/abs/2202.05786

A 相关工作

A.1 实体对齐

实体对齐（EA）技术主要包括早期的基于嵌入的方法 [1, 6] 和近年来 [2, 17] 流行的基于图神经网络的方法。各种基于图神经网络的方法主要关注属性 [2, 15]、关系 [5, 24] 和 [20, 26] 邻居特征的聚合方式。具体而言，考虑属性的方法会聚合多种类型的属性，[2, 5, 31] 或结合多个模型对实体进行编码，从多个角度学习实体嵌入。AttrGNN [15] 将知识图谱划分为多个子图以聚合属性，从而有效地对各种类型的属性三元组进行建模。上述方法证明了聚合属性对实体对齐的有效性。然而，这些方法都忽略了属性的不一致性以及图像属性。

A.2 多模态实体对齐

此外，由于现实世界中的知识图谱具有多模态特性，已有若干研究 [10, 22, 33] 开始关注多模态实体对齐技术。与单模态知识图谱上的实体对齐类似，多模态知识图谱中的许多任务 [13, 19] 为融合多模态属性和关系提供了可能。作为首个多模态实体对齐研究，PoE [14] 将每个实体表示为一个单一向量，其中拼接了实体的所有模态特征。然而，该方法无法捕捉异构模态之间潜在的交互作用，限制了其准确执行实体对齐的能力。随后，Chen 等人 [3] 提出了一种多模态知识嵌入方法，用于区分性地生成不同类型知识的知识表示，并设计了一个多模态融合模块将其整合。Guo 等人 [7] 提出了双曲多模态实体对齐（HEA）方法，在双曲空间中结合属性表示和实体表示，并利用聚合嵌入预测对齐。MCLEA [12]和 MSNEA [4] 缩小了每个实体不同模态之间的差异，并利用名称嵌入。然而，现有方法忽略了实体对之间的上下文差异，进而可能削弱对齐效果。

B 数据集与评估指标

B.1 数据集

在我们的实验中，我们使用了两个基于 [14], 构建的多模态数据集，即 FB15K-DB15K 和 FB15K-YAGO15K。FB15K 是从 Freebase 知识库中提取的一个具有代表性的子集。为了保持与 FB15K 近似的实体数量，DBpedia 中的 DB15K 和 YAGO 中的 YAGO15K 主要根据与 FB15K 对齐的实体进行选取。表 3 展示了多模态数据集的统计信息。

表 3：数据集统计信息。（Rel.：关系，Attr.：属性，Rel. T.：关系三元组，Attr. T.：属性三元组。）

数据集	#实体	#关系	属性	关系类型	属性类型	图像
FB15K	14,951	1,345	116	592,213	29,395	13,444
DB15K	12,842	279	225	89,197	48,080	12,837
YAGO15K	15,404	32	7	122,886	23,532	11,194

B.2 评估指标

我们利用余弦相似度计算两个实体之间的相似度，并采用 Hits@n 和平均倒数排名作为评估全部模型的指标。Hits@n 表示根据相似度计算结果，正确实体排名位于前 n 位的比例。平均倒数排名表示正确实体的平均倒数排名。Hits@n 和平均倒数排名的值越高，说明该方法的性能越好。

C 超参数

为便于复现并促进研究，我们在表4中报告了超参数设置。请注意，所有超参数设置均通过网格搜索在验证集上进行调优，共进行5次试验。我们采用 Hugging Face 中的 bert-large-uncased 作为编码器，其层数为24，嵌入大小为1024。超参数 λ_1，λ_2，λ_3 的最佳取值分别为5、3和2。具体而言，如果模型连续100轮未能降低验证集上的损失函数值，则停止运行。所有基线模型均使用相同的数据集划分，以确保公平性。为确保公平，所有基线方法均使用相同的实体表示维度，设为128维。所有实验均在配备一块图形处理器（Tesla V100）的服务器上进行。

D 实证运行时分析

所提出框架的时间复杂度可以接受。表 5 展示了我们的方法与三种性能良好的基线方法在 FB15K-DB15K 和 FB15K-YAGO15K 上的训练耗时。为确保公平，我们只使用一个图形处理器，包括 AttrGNN。因此，我们依次训练 AttrGNN 的三个通道，这使得该模型的训练时间最长。我们的方法会生成属性一致性知识图谱，从而增加时间开销。不过，我们设计了随机丢弃机制来节省时间。总体而言，我们方法的时间复杂度可与其他高效方法相媲美。

表4：模型的超参数设置。

超参数	**FB15K-DB15K**	**FB15K-YAGO15K**
批量大小	512	512
训练轮次	200	200
学习率	0.001	0.001
温度	0.5	0.5
负样本数量	15	15
权重衰减	0.01	0.01
随机丢弃率	0.35	0.35
λ_1	5	5
λ_2	3	3
λ_3	2	2

表5：FB15K-DB15K 和 FB15K-YAGO15K 数据集上的平均训练时间（秒）。

方法	AttrGNN	PoE	PoE-rni	ACK-MMEA
FB15K-DB15K	396	164	162	165
FB15K-YAGO15K	389	155	151	156