# 思维交流与写作

实践四：论文写作模拟

毕业设计论文一般包括摘要和五个章节内容，第一章节和第五章节的题目分别为引言和结论，第二章节、第三章节和第四章节的题目自拟。

任务一：完成毕业设计选题后，以周为单位简要制定每周毕业设计论文进度表，总计16周。（见附录）okkk

任务二：用自己话描述毕业设计论文中【中文摘要】、【引言】、【结论】的作用与应包含的内容，每部分字数200字左右。Ok

任务三：用自己的话描述毕业设计论文中第二章节、第三章节和第四章节应包括的内容，每部分字数200字左右。

任务四:填写毕业设计模板封面学院、专业、学生姓名、学号、指导教师(任课教师):将实践二中的参考文献按要求填写入参考文献处。Ok

所有任务都在附件文件中完成，并上传。

附 录

附录A 毕业设计（论文）进度安排：

|  |  |  |  |
| --- | --- | --- | --- |
| 毕业设计（论文）进度安排： | | | |
| 序号 | 毕业设计（论文）各阶段内容 | 时间安排 | 备注 |
| 1 | 深入理解课题，精读核心文献，收集并筛选相关背景资料（SGG, 原型学习等）。 | 第一周 | 提交开题报告，明确要解决的核心问题。 |
| 2 | 广泛阅读相关文献，完成文献综述的撰写，并初步拟定论文提纲。 | 第二周 | 重点理解SGG领域的常用基线模型和长尾问题解决方案。 |
| 3 | 学习并配置相关实验环境，包括PyTorch、CUDA等，下载并预处理Visual Genome数据集。 | 第三周 | 需提前准备 |
| 4 | 搭建并运行一个基础的SGG基线模型，复现其基本结果，熟悉代码框架。 | 第四周 | 确保代码能跑通，为后续模型改进打下基础。 |
| 5 | 基于原型学习理论，初步设计一个全新的SGG模型框架。构思如何将原型与视觉、空间信息结合。 | 第五周 | 撰写详细的研究方案，包括创新的技术构想和可行性分析。 |
| 6 | 基于上周的设计，编码实现模型的核心表征模块，即如何为每个实例动态生成融合了原型和上下文的特征。 | 第六周 | 理论转化实践 |
| 7 | 探索并设计一种机制来增强同类谓词表征的相似度。构思使用何种几何约束 | 第七周 | 进行相关实验 |
| 8 | 将设计的几何约束机制编码并集成到模型中，进行初步的模块功能验证。 | 第八周 | 观察实验结果 |
| 9 | 为解决谓词混淆问题，构思一种能主动推开不同谓词原型的正则化方法。 | 第九周 | 进行相关实验 |
| 10 | 运行完整的模型，与基线模型进行初步对比，验证所提方法的整体有效性。 | 第十周 | 证明方法有效性 |
| 11 | 模型组件设计并进行消融实验，分别验证创新模块对最终性能的贡献。 | 第十一周 | 消融实验 |
| 12 | 结果整理与论文撰写，全面整理所有实验数据，绘制图表，并完成论文初稿。 | 第十二周 | 论文初稿 |
| 13 | 导师反馈与论文修改，将论文初稿提交给指导老师审阅，根据反馈意见进行深入修改和完善。 | 第十三周 | 反馈给导师 |
| 14 | 补充实验与答辩准备，根据修改意见，有选择地补充少量实验。开始构思并制作答辩PPT。 | 第十四周 | 准备答辩材料 |
| 15 | 论文终稿与格式审查，完成论文的全部修改，进行最终的语言润色、排版和格式检查。 | 第十五周 | 论文最终审查 |
| 16 | 提交终稿与答辩预演，提交最终版论文，并进行答辩的模拟演练，熟悉流程，准备应对各种提问。 | 第十六周 | 模拟答辩 |



本科毕业设计（论文）

**基于原型关系约束的鲁棒场景图生成方法研究**

**Robust Scene Graph Generation**

**Based on Prototype Relationship Constraints**

**（不必删除，下方开始完成《思维交流与写作》课程实践）**

学 院： 计算机科学与技术学院

专 业： 计算机科学与技术

学生姓名： 江家玮

学 号： 22281188

指导教师： 秦颖

**北京交通大学**

2025年7月

学士论文版权使用授权书

本学士论文作者完全了解北京交通大学有关保留、使用学士论文的规定。特授权北京交通大学可以将学士论文的全部或部分内容编入有关数据库进行检索，提供阅览服务，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。

（保密的学位论文在解密后适用本授权说明）

学位论文作者签名： 指导教师签名：

签字日期： 年 月 日 签字日期： 年 月 日

中文摘要

**摘要：**场景图生成（SGG）是一项富有挑战性的模式识别任务，其核心在于识别图像中的物体及其成对关系。当前SGG方法的一个关键瓶颈是为关系模式（谓词）学习具有区分度的表征，这一难题因显著的类内视觉差异和类间相似性而加剧，导致识别模糊和性能偏差，尤其对低频模式影响更大。为解决这些表征学习的挑战，本文提出了一种基于原型关系约束的模块(PRCon-PM)。该方法利用原型学习理论…….为了增强表征的区分能力，我们引入了两种关键机制……在基准数据集上的大量实验表明，PRCon-PM显著提升了谓词识别的准确性和鲁棒性，取得了业界领先的性能，尤其在正确分类低频关系模式方面效果显著。

**关键词：**场景图生成；关系推理；注意力机制；上下文信息

**以大创选题，鲁棒场景图生成方法的设计与实现为例, 我进行了上述摘要的部分书写。**我认为摘要先对当前场景图生成的意义和现状做一个简要的阐述(研究目的和基本研究内容)，而后开始阐述目前场景图生成中所遇到的问题和挑战，之后基于上述的问题和挑战，我提出了自己的某一个方法去进行解决，然后重点突出，语言精炼的模块阐述我的方法，之后在写我的方法所带来的结果提升，以证明我的方法有效性(研究方法、结果和结论)。在这个过程中，要注意突出论文的创新点。同时我认为，摘要需要有独立性和自明性，即不用阅读全文，就能获得论文必要的信息。并且摘要中不应该使用公式、图表，以及引用文献。

**（不必删除，下方开始完成《思维交流与写作》课程实践）**

ABSTRACT

**ABSTRACT:** Scene Graph Generation (SGG) is a challenging pattern recognition task for identifying objects and their pairwise relationships within images. A critical bottleneck is learning discriminative representations for relationship patterns (predicates), which is hampered by significant intra-class visual variations and inter-class similarities. These factors cause ambiguity and biased recognition performance, particularly for infrequent patterns. To address these representational challenges, this paper proposes the Prototype-based Relation Constrained Module (PRCon-PM). Our approach leverages prototype learning theory… To enhance discriminative power, we introduce two key mechanisms… Extensive experiments on benchmark datasets demonstrate that PRCon-PM significantly improves the accuracy and robustness of predicate recognition, yielding state-of-the-art performance, especially in correctly classifying less frequent relationship patterns 。

**KEYWORDS：**Scene Graph Generation; Relation Inference; Attention Mechanism; Contextual Information

目 录

[中文摘要 i](#_Toc414268225)

[ABSTRACT ii](#_Toc414268226)

[目 录 iii](#_Toc414268227)

[1 引言 1](#_Toc414268228)

[2 （小四、黑体、左对齐、空1字符） 2](#_Toc414268229)

[2.1 （小四、宋体、首行缩进1字符） 3](#_Toc414268230)

[2.1.1 （小四、宋体、首行缩进2字符） 3](#_Toc414268231)

[3 4](#_Toc414268232)

[3.1 4](#_Toc414268233)

[3.1.1 4](#_Toc414268234)

[4 5](#_Toc414268235)

[4.1 5](#_Toc414268236)

[4.1.1 5](#_Toc414268237)

[5 结论 6](#_Toc414268238)

[参考文献 7](#_Toc414268239)

[致 谢 8](#_Toc414268240)

[附 录 8](#_Toc414268240)

目录说明：

（小四、黑体、左对齐、空1字符）

（小四、宋体、首行缩进1字符）

（小四、宋体、首行缩进2字符）

1. 引言

引言（或绪论）简要说明研究工作的目的、范围、相关领域的前人工作和知识空白、理论基础和分析、研究设想、研究方法和实验设计、预期结果和意义等。应言简意赅，不要与摘要雷同，不要成为摘要的注释。一般教科书中有的知识，在引言中不必赘述。

**（不必删除，下方开始完成《思维交流与写作》课程实践）**

**引言的写作思路：**

**1. 研究背景：**首先从一个宏观的视角切入，介绍我这篇文章的研究领域。解释这个领域为什么重要，并且解决了什么样的大问题。让一个非本领域的专家也能大致明白我的研究是关于什么的。

**2. 研究现状与挑战:** 将宏观背景聚焦要解决的那个具体问题上。清晰地定义要研究的具体任务。其实就是要阐述我的研究动机，找到挑战，也就是paper里面常说的challenges，大概可以写为：尽管已经有很多研究，但仍然存在哪些尚未解决的问题、挑战或局限性。

**3. 本文的研究工作：**直接、清晰地说明采用了什么方法或模型来解决上述问题。

【接下来以我大创的项目为例进行书写】

* 1. 研究意义与背景

随着人工智能和计算机视觉技术的飞速发展，让机器从“看见”图像到“理解”图像中的复杂场景，已成为该领域的核心目标之一。仅仅识别出图像中的独立物体是远远不够的，更关键的是要理解这些物体之间丰富的交互和语义关系。在这一背景下，场景图生成（Scene Graph Generation, SGG）作为一项关键技术应运而生。SGG旨在将图像解析为一个结构化的图谱，其中节点代表图像中的物体实例，有向边则代表物体间的成对关系（例如，<人, 骑在, 马>）。这种富含语义的结构化表征，为众多下游视觉任务，如视觉问答（VQA）、图像描述和图像检索 等提供了强大的支持，因此具有重要的研究价值和广阔的应用前景。

* 1. 研究现状与挑战

尽管SGG领域已取得显著进展，但生成准确且鲁棒的场景图仍面临巨大挑战。现有方法通常采用物体检测器定位物体，再利用上下文建模模块（如消息传递网络 或循环神经网络）来预测关系。然而，一个核心的瓶颈在于学习具有高区分度的关系谓词表征。这一挑战主要源于两个方面：

第一，巨大的类内视觉差异。同一个语义关系，在不同场景下其视觉外观可能千差万别。例如，“吃”这一关系可以表现为“人吃披萨”，也可以是“长颈鹿吃树叶”，二者视觉特征差异悬殊。

第二，严重的类间视觉相似性。不同的关系可能在视觉上极为相似，难以区分。例如，“男人拿着盘子”和“男人吃着披萨”在视觉上就十分接近。这两个问题共同导致了模型在谓词识别上的模糊性，限制了其泛化能力。此外，真实世界的数据集（如Visual Genome）普遍存在长尾分布问题，模型训练后会偏向于预测高频、但信息量可能较低的关系（如 'on', 'has'），而难以识别稀有但语义更丰富的关系，这严重影响了场景图的质量和实用性。

* 1. 研究难点

为了克服上述在关系表征学习中的根本性限制，探索一种能够学习更鲁棒、更具区分性的谓词表征的场景图生成方法。本文提出xxxx方法，该方法的核心思想为xxxx

1. 相关技术与基准模型

**（不必删除，下方开始完成《思维交流与写作》课程实践）**

这个章节作为一个承上启下的作用，承接引言中提出的问题，通过介绍一个具体的、有代表性的现有模型（Motifs），将宏观的问题“具象化”。主要可以看到现有技术是如何工作的，与具体的短板在哪里。

并且为我下一章节所提出的新的模型提供了参考。

场景图生成（SGG）任务旨在将图像内容解析为一个结构化的图，其中为物体集合，为关系集合。每个关系通常由一个<主体-谓词-客体>三元组构成。尽管现有方法众多，但大多遵循一个相似的两阶段流程：首先利用物体检测器识别物体，再通过一个关系预测模块判断物体间的关系。为客观评估本研究提出方法的有效性，选取了经典的Neural Motifs (Motifs) 模型….

学位论文为了需要反映出作者确已掌握了坚实的基础理论和系统的专门知识，具有开阔的科学视野，对研究方案作了充分论证，因此，有关历史回顾和前人工作的综合评述，以及理论分析等，可以单独成章，用足够的文字叙述。正文是学位论文的核心部分，占主要篇幅，可以包括：调查对象、实验和观测方法、仪器设备、材料原料、实验和观测结果、计算方法和编程原理、数据资料、经过加工整理的图表、形成的论点和导出的结论等。

由于研究工作涉及的学科、选题、研究方法、工作进程、结果表达方式等有很大的差异，对正文内容不能作统一的规定。但是，必须实事求是，客观真切，准确完备，合乎逻辑，层次分明，简练可读。

**图：**

图包括曲线图、构造图、示意图、框图、流程图、记录图、地图、照片等。

图应具有“自明性”。

图应有编号。图的编号由“图”和从“1”开始的阿拉伯数字组成，图较多时，可分章编号。

图宜有图题，图题即图的名称，置于图的编号之后。图的编号和图题应置于图下方。

照片图要求主题和主要显示部分的轮廓鲜明，便于制版。如用放大缩小的复制品，必须清晰，反差适中。照片上应有表示目的物尺寸的标度。

图片示例1：

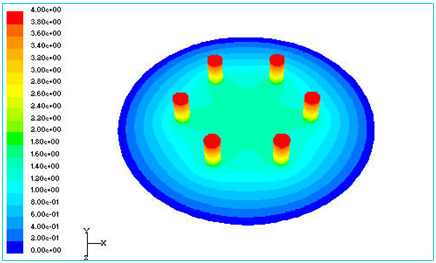


图2-1 太合金多炭钢铁产品柱扭曲局部受力分析示意图

**表：**

表应具有“自明性”。

表应有编号。表的编号由“表”和从“1”开始的阿拉伯数字组成，表较多时，可分章编号。

表宜有表题，表题即表的名称，置于表的编号之后。表的编号和表题应置于表上方。

表的编排，一般是内容和测试项目由左至右横读，数据依序竖读。

表的编排建议采用国际通行的三线表。

如某个表需要转页接排，在随后的各页上应重复表的编号。编号后跟表题（可省略）和“（续）”，置于表上方。

续表均应重复表头。

表格示例1：

表2-1 国际单位制的基本单位

|  |  |  |
| --- | --- | --- |
| 量的名称 | 单位名称 | 单位符号 |
| 长度 | 米 | m |
| 质量 | 千克(公斤) | kg |
| 时间 | 秒 | s |
| 电流 | 安[培] | A |
| 热力学温度 | 开[尔文] | K |
| 物质的量 | 摩[尔] | mol |
| 发光强度 | 坎[德拉] | cd |

**公式：**

论文中的公式应另行起，并缩格书写，与周围文字留足够的空间区分开。

如有两个以上的公式，应用从“1”开始的阿拉伯数字进行编号，并将编号置于括号内。公式的编号右端对齐，公式与编号之间可用“…”连接。公式较多时，可分章编号。

公式示例1：

 (2-1)

 (2-2)

式中 —— 多孔质材料的平均粒子直径(m)；

—— 孔隙度（孔隙体积占总体积的百分比）；

—— 特征渗透性或固有渗透性，与材料的结构性质有关(m2)。

较长的公式需要转行时，应尽可能在“＝”处回行，或者在“+”、“－”“×”、“/”等记号处回行。

公式中分数线的横线，其长度应等于或略大于分子和分母中较长的一方。

如正文中书写分数，应尽量将其高度降低为一行。如将分数线书写为“/”，将根号改为负指数。

公式示例2：

将  写成 1/ 或 

**引文标注**

论文中引用的文献的标注方法遵照GB/T 7714－2005，可采用顺序编码制，也可采用著者－出版年制，但全文必须统一。

**注释**

当论文中的字、词或短语，需要进一步加以说明，而又没有具体的文献来源时，用注释。注释一般在社会科学中用得较多。

应控制论文中的注释数量，不宜过多。

由于论文篇幅较长，建议采用文中编号加“脚注”的方式。最好不用采用文中编号加“尾注”。

1. 基于原型关系约束的场景图生成模型

**（不必删除，下方开始完成《思维交流与写作》课程实践）**

为解决传统方法在谓词表征学习上的根本性缺陷，我们提出一种全新的基于原型关系约束的场景图生成模型。其核心思想是，不再仅仅依赖于不稳定的视觉特征，而是将关系表征锚定在更稳定、更具区分性的语义“原型”上。我们首先为每个物体和谓词类别（如“人”、“马”、“骑”）在语义空间中定义一个可学习的原型向量……

在这个章节，我将开始具体的阐述我的方法，然后把我提出的核心的模型的总体设计的思想进行书写。

1. 实验设计与评测标准

**（不必删除，下方开始完成《思维交流与写作》课程实践）**

为科学、全面地验证本研究提出模型的有效性，设计了严谨的实验流程。实验将基于两个业界公认的大规模数据集：Visual Genome……

在这个章节中，我将主要讲我的实验设计，具体的实施和相关的参数，目的是表明我的方法的有效性，实验结果是可信、可复现的。这个是衡量我工作的一个重要的标准，其中会将我的结果和近期相关工作以及基线进行对比，证明有效性。然后会进行消融实验，保证我设计的模块是有效的，包括一些参数设置的实验等等。

这个章节需要包括数据集的描述，评测的任务等，所实现的细节，实验的方案等等。

1. 结论与展望

论文的结论是最终的、总体的结论，不是正文中各段的小结的简单重复。结论应该准确、完整、明确、精练。如果不可能导出应有的结论，也可以没有结论而进行必要的讨论。可以在结论或讨论中提出建议、研究设想、仪器设备改进意见以及尚待解决的问题等。

**（不必删除，下方开始完成《思维交流与写作》课程实践）**

**我认为结论不是内容的简单重复，是对整个研究工作的提炼、升华和反思。我个人认为结论的作用是：让读者在读完后，对我的工作有一个完整、清晰且深刻的印象。因此总结流程，我认为结论的书写大致是：总结工作(重申研究问题+概括方法+再次说我的贡献) → 分析局限(客观的列举出我的不足之处) → 展望未来(基于局限性，提出未来的改进方向)**

**接下来我继续基于我大创场景图生成进行书写**

* 1. 本文工作总结

本文主要围绕计算机视觉领域中的一个关键而富有挑战性的任务——场景图生成（Scene Graph Generation, SGG）展开研究。具体而言，本文聚焦于SGG中的一个核心瓶颈：由于关系谓词存在巨大的类内视觉差异和严重的类间视觉相似性，导致模型难以学习到鲁棒且具有高区分度的谓词表征。

为应对这一挑战，本文深入研究并实现了一种基于原型关系约束的场景图生成方法（PRCon-PM）。该方法的核心是xxxx

* 1. 局限性分析

尽管本文研究的方法取得了良好的效果，但仍存在一些局限性：xxxx

* 1. 未来工作展望

基于本文的研究工作和存在的局限性，未来的研究可以从以下几个方面展开：

参考文献

参考文献是文中引用的有具体文字来源的文献集合。按照GB 7714《文后参考文献著录规则》的规定执行。

参考文献以文献在整个论文中出现的次序用[1]、[2]、[3]……形式统一排序、依次列出。

**著作:**

[序号] 作者.译者.书名.版本.出版地.出版社.出版时间.引用部分起止页

**期刊:**

[序号] 作者.译者.文章题目.期刊名.年份.卷号(期数).引用部分起止页

**会议论文集：**

[序号]作者.译者.文章名.文集名 .会址.开会年.出版地.出版者.出版时间.引用部分起止页

**学位论文：**

[序号]作者.题名[学位论文]（英文用[Dissertation]）.保存地点.保存单位.年份.引用部分起止页

**专利:**

[序号] 专利申请者.题名.国别.专利文献种类.专利号.发布日期.引用部分起止页

**技术标准:**

[序号] 起草责任者.标准代号.标准顺序号—发布年.标准名称.出版地.出版者.出版年份.引用部分起止页

**（不必删除，下方开始完成《思维交流与写作》课程实践）**

1. Wu S, Chen N, Xiao A, et al. AI-Empowered Virtual Network Embedding: A Comprehensive Survey[J]. IEEE Communications Surveys and Tutorials, 2025, 27(2): 1395-1426.
2. Petrenko S, Da Silva L E B, Wunsch Ii D C. DeepART: Deep gradient-free local learning with adaptive resonance[J]. Neural Networks, 2025, 190: 107615.
3. Lin X, Wang S, Cai R, et al. Suppress and rebalance: Towards generalized multi-modal face anti-spoofing[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2024: 211-221.
4. Ai H, Wang L. Elite360d: Towards efficient 360 depth estimation via semantic-and distance-aware bi-projection fusion[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2024: 9926-9935.
5. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).Las Vegas, NV, USA: IEEE, 2016: 770-778.
6. Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detectionwith Region Proposal Networks[C]//Advances in Neural Information Processing Systems (NIPS). Montreal, Canada: Curran Associates, Inc., 2015: 91-99.