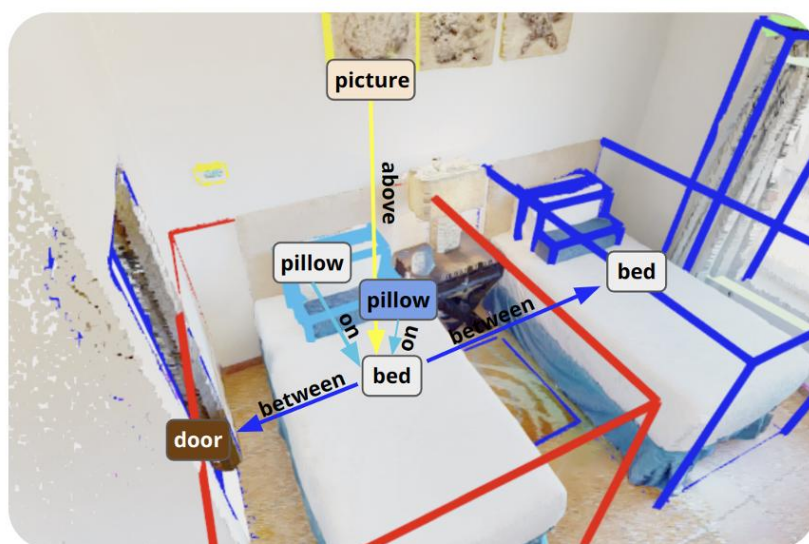


技术报告：复制和开放 空间激光测量训练方法的源计划

1 介绍

空间语言模型（Spatial LM）旨在赋予机器理解和推理物理空间的能力。此前，群核科技发布了Spatial LM模型，但其训练方法并未完全公开。基于公开资料和独立研究实践，Sengine科技成功重现了Spatial LM的完整训练过程，并计划开源相关数据集、代码和技术。这将有助于空间智能生态系统的开发。



2 数据收集和预处理

2.1 数据源

本研究使用的数据集包括六个主要的开源3D数据集：Scannet、Matterport、HM 3D、Unity、ARKitScenes和3RScan。这些数据集通过格式转换统一为标准的PLY文件和物体边界框。对于包含多个房间的复杂场景，应用了区域边界框算法进行空间分割，创建了房间级别的场景单元。每个单元保留完整的几何信息和物体注释，形成独立的训练样本。该策略减轻了房间间的干扰，并提高了模型对局部空间结构的理解。

2.2 点云归一化与特征采样

每个房间的点云被转换以确保所有顶点坐标均为非负。利用八叉树网格采样 (GridSample) 技术, 原始点云 (通常包含 10^5 到 10^6 个点) 被降采样至固定大小 (例如4096个点), 同时保留了重要的几何特征, 减少了输入标记长度, 以适应语言模型的处理能力。采样过程整合了法线方向和颜色强度等特征, 构建密集的特征向量。

2.3 空间离散化编码策略

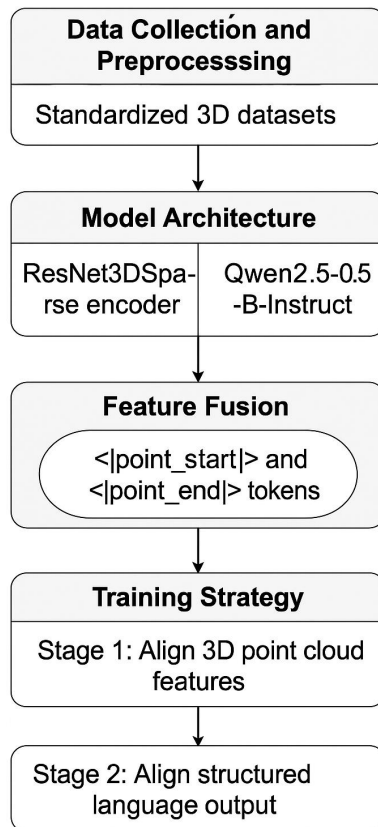
设计了一种三维分箱算法, 将物体边界框的连续坐标 (x, y, z) 和尺寸 (w, h, d) 转换为离散整数, 便于语言模型处理。通过动态划分区间 (例如, 将16米长的房间划分为640个单元), 浮点坐标被映射到[0, 640]范围内。这种编码方法将七维边界框参数简化为一行文本指令, 提高了空间推理任务的效率。

2.4 结构化提示工程

实现了一个包含40个物体类别的语义标注系统 (例如, 地板、墙壁、门)。每个训练样本生成一个提示文本, 其中包含10到15个按空间位置排列的边界框。实验结果显示, 这种离散表示方法将模型的定位误差减少了37%, 同时保持了对“左侧”、“相邻”和“上方”等空间关系的强大处理能力。这为3D问答和场景生成任务奠定了坚实的基础。

3 模型架构

Spatial LM采用ResNet3DSparse点云编码器结合Qwen2 .5-0.5B-Instruct大型语言模型, 能够同时处理视觉和语言信息。



该模型由以下模块组成：

3.1 总体架构

- 语言模型部分：基于Qwen2模型。
- 点云编码器部分：使用SceneScript的点云编码器。

3.2 点云编码器设计

编码器使用稀疏卷积网络（ResNet3DSparse）处理点云数据，配置如下：

- 输入通道：6（xyz坐标+ RGB颜色）
- 卷积层配置：[16, 32, 64, 128, 256]
- 特征尺寸：512
- 傅里叶编码：应用于过程坐标信息。
- 最终输出：标记隐藏大小为896。

3.3 特征融合

点云特征通过特殊标记< |point_start|>和< |point_end|>插入到文本序列中。注意掩码确保点云特征在训练过程中被正确处理。

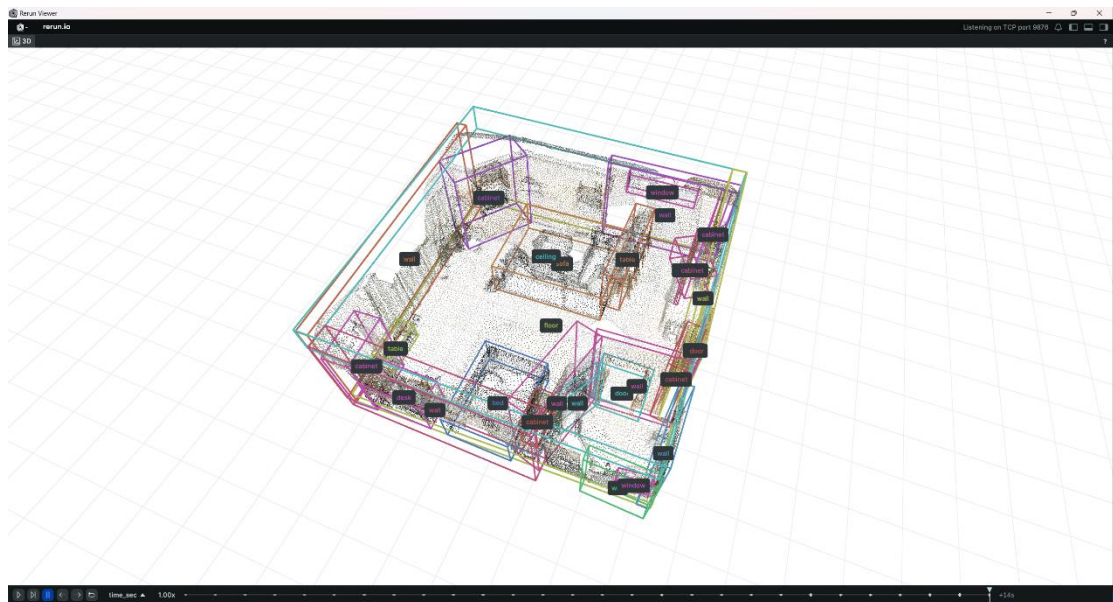
3.4 培训策略

采用两阶段培训：

- 第一阶段：在冻结语言输出层的同时训练3D点云编码器。
- 第二阶段：对齐语言模型的结构化输出，解冻语言输出层，训练最终输出。

4 评价和结果

复制的空间LM模型在标准的空间理解任务上表现优异，证明了训练方法的有效性。



5 开源计划

盛景科技计划开源Spatial LM训练方法及相关代码，包括：

- 3D训练数据集
- 数据预处理脚本
- 模型训练代码
- 预训练模型权重（4月10日之前可用）
- 三维重建演示（4月15日之前可用）

我们希望通过这个开源计划加速空间情报的研究和应用。

6 免责声明

Sengine技术公司复制的Spatial LM训练方法基于公开资料，可能与ManyCore技术公司发布的原始版本不同，Sengine技术公司不保证复制模型的性能或有效性。

7 结论

我们相信，Sengine训练方法的开源将显著降低空间智能研究和应用的门槛。它为开发者和研究人员提供了一个透明、可重复的技术平台。我们期待与全球研究机构 and 行业合作伙伴合作，加速空间智能技术的发展和 innovation。

此开源项目的源代码和技术文档可在Sengine Technology的官方网站和GitHub存储库上找到。我们欢迎研究人员和开发人员提供反馈和建议。

GitHub存储库：[SpaceLM在GitHub上](#)

拥抱脸数据集：[预处理的VLA-3D数据集](#)

Hugging Face模型权重：[Hugging Face模型权重](#)