目 录

[0 引言 1](#_Toc198989859)

[1 相关工作概述 1](#_Toc198989860)

[1.1 城市流量数据预测 1](#_Toc198989861)

[1.2 城市流量数据超分辨率 1](#_Toc198989862)

[2 基于时空吸引学习的不可观测数据下细粒度城市流量推理 1](#_Toc198989863)

[2.1 引言 1](#_Toc198989864)

[2.2 方法描述 2](#_Toc198989865)

[2.2.1 符号说明与问题定义 2](#_Toc198989866)

[2.2.2 城市流量数据补全网络 3](#_Toc198989867)

[2.2.3 城市流量数据补全网络 5](#_Toc198989868)

[2.2.4 最终目标函数 5](#_Toc198989869)

[2.3 实验与分析 6](#_Toc198989870)

[2.3.1 实验设置 6](#_Toc198989871)

[2.3.1.1 数据集 6](#_Toc198989872)

[2.3.1.2 对比方法 6](#_Toc198989873)

[2.3.1.3 训练细节与超参数 6](#_Toc198989874)

[2.3.2 与基线方法的对比分析 6](#_Toc198989875)

[2.3.3 消融分析 9](#_Toc198989876)

[2.3.4 参数分析 9](#_Toc198989877)

[4 结论（结论和建议） 9](#_Toc198989878)

[参考文献 10](#_Toc198989879)

插图清单

[图 1 两种粒度层次的城市流量图示意图 2](#_Toc198989880)

[图 2 UrbanSTA 框架结构图。该方法由两个主要部分组成：时空注意力补全网络（STA）与细粒度推理网络（FIN）。STA 包含非对称的编码器与解码器，FIN 仅包含解码器。UrbanSTA 采用三阶段结构，将多任务问题分解为时空特征编码与解码步骤，以实现流量补全与超分辨率推理。 4](#_Toc198989881)

[图 3 （a）STA 模块在任务中的有效性分析；（b）FIN 模块在任务中的有效性分析。 7](#_Toc198989882)

[图 4 在 TaxiBJ 数据集（缺失率 20%，上采样因子 4）上的不同超参数设置下模型性能对比。 8](#_Toc198989883)

附表清单

[表 1 数据集统计信息 6](#_Toc198989884)

[表 2 TaxiBJ 数据集上的对比实验结果。2 和 4 表示上采样因子，20%、40%、60% 表示数据缺失率。最佳结果加粗，次佳结果下划线标记。 7](#_Toc198989885)

[表 3 BikeNYC 数据集上的对比实验结果。2 和 4 表示上采样因子，20%、40%、60% 表示数据缺失率。最佳结果加粗，次佳结果下划线标记。 8](#_Toc198989886)

[表 4 消融实验结果。在 TaxiBJ 数据集（缺失率 20%，上采样因子 4）上报告不同约束条件下模型性能。最佳结果加粗，次佳结果下划线标记。 9](#_Toc198989887)

# 0 引言

# 1 相关工作概述

## 1.1 城市流量数据预测

城市流量数据预测在城市计算领域中占据重要地位，研究方法多种多样[9,10,11]。在深度学习方法中，DeepST 模型[12]被提出，用于建模时间与空间特征，并能同时提取外部因素的影响。另一典型模型 ST-ResNet [13]则通过残差单元与特征融合组件以学习更复杂的时空特征。

与传统的城市流量数据预测不同，本文研究的是在粗粒度观测数据存在缺失情况下，推理细粒度城市流量的任务。

## 1.2 城市流量数据超分辨率

深度学习方法是当前推理细粒度城市流量的主流手段[14]。UrbanFM 模型[15]通过引入外部因子融合网络以提取外部特征，从而预测城市交通流量。为进一步提升 UrbanFM 的推理性能，Ouyang 等人[16]提出了 UrbanPy 模型，分为两个部分：推理网络用于生成细粒度流量分布，通用融合子网络用于增强整体性能。

然而，上述研究多基于理想数据，现实中普遍存在缺失问题。Li 等[6]首次针对该问题提出了 MT-CSR 多任务网络模型，设计了 CMPNet 补全网络，综合考虑局部空间依赖性与全局 POI 相似性以填补粗粒度流量图的缺失；此外还提出了 SRNet 超分辨率网络，用以捕捉粗细粒度数据之间的复杂关联。

# 2 基于时空吸引学习的不可观测数据下细粒度城市流量推理

## 2.1 引言

随着智慧城市建设的加速推进[1]，城市流量预测成为城市化与智慧城市发展的关键组成部分。高精度的城市流量预测对城市管理者具有重要意义，有助于未来智慧城市的持续建设[2]。然而，仅通过部署传感器获取城市流量数据在实际中往往难以满足需求[3,4,5]，受制于实施成本、地域限制等因素，部分区域流量数据缺失成为普遍现象[6]。

当前关于细粒度城市流量推理的研究仍相对有限[7,8]，且多集中于将缺失值预测与细粒度推理视作两个独立任务，而忽略了现实中两者往往需要同时处理[6]。为此，本文提出了 UrbanSTA 模型，旨在解决以下三个主要挑战：

1. 城市流量图的不完整性：由于传感器分布不均和区域覆盖受限，导致流量图中存在缺失值；
2. 粗细粒度流量之间的复杂相关性：一个粗粒度区域通常对应多个相邻细粒度区域，推理难度大；
3. 城市流量图中空间属性的相互依赖性：区域流量通常受其邻域及相似区域流量变化的影响。

UrbanSTA 包含两个子网络：STA 模块采用非对称编码-解码结构，通过提取空间与时间特征来预测缺失值；FIN 模块则在空间注意力机制下利用预测出的粗粒度流量推断细粒度流量，同时采用分布式上采样机制以恢复原始结构约束关系。本文的主要贡献如下：

1. 城市流量图中空间属性的相互依赖性：区域流量通常受其邻域及相似区域流量变化的影响。
2. 设计分布式上采样策略以建模粗细粒度城市流量之间的复杂关系；
3. 考虑空间旋转不变性与时间周期性，引入时空吸引约束损失以增强时空表征能力。

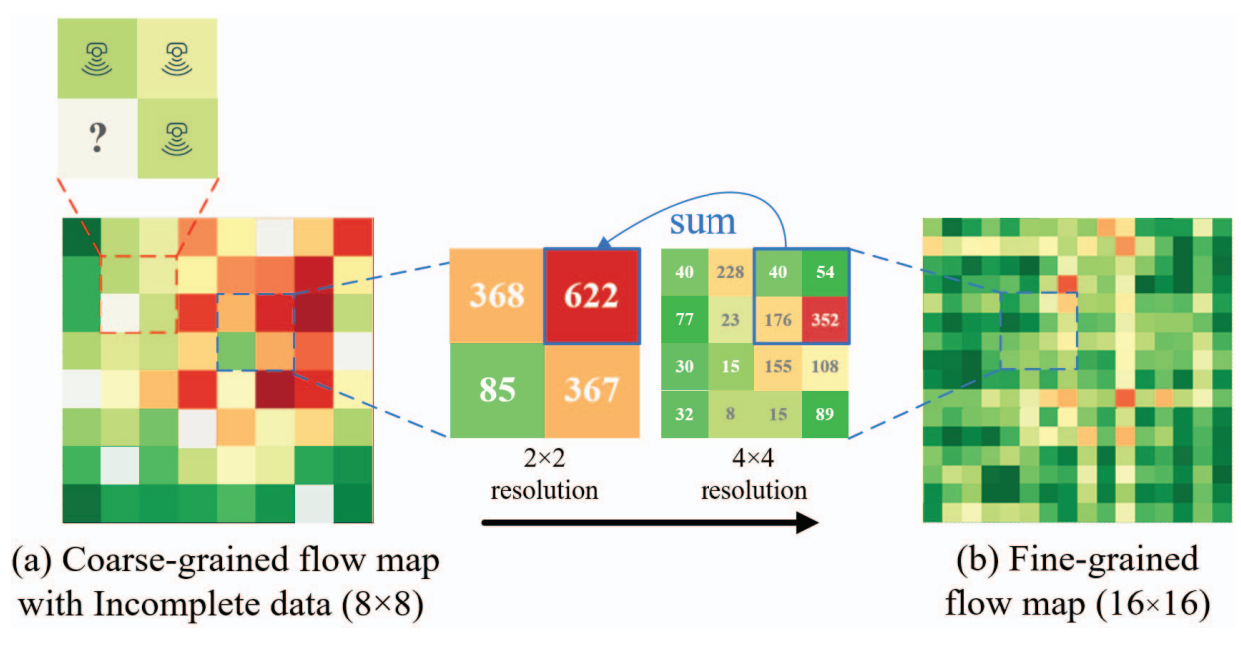
****

图 1 两种粒度层次的城市流量图示意图

## 2.2 方法描述

本文提出的 UrbanSTA 模型如图 2 所示，旨在处理不可观测数据下的细粒度城市流量推理问题。

### 2.2.1 符号说明与问题定义

#### 2.2.1.1 定义1：粗粒度与细粒度城市流量图

一般而言，原始传感器所采集的 数据被定义为粗粒度城市流量图 ，通过推理可得细粒度城市流量图 ，其中 为上采样因子。

#### 2.2.1.2 定义2：不完整的粗粒度城市流量图

在现实中，由于某些区域长期缺失观测数据，粗粒度城市流量图呈现出固定位置的缺失模式，记作。

#### 2.2.1.3 定义3：结构约束

每个粗粒度区域的流量 等于其对应细粒度区域集合的流量总和，即：

其中满足，。

#### 2.2.1.4 目标定义

给定上采样因子 和一组历史及当前的不完整粗粒度流量图 ，本研究的目标是推理当前完整的细粒度城市流量图 。

### 2.2.2 城市流量数据补全网络

#### 2.2.2.1 初始化

**掩码**。为了对缺失区域进行城市流量数据补全，我们定义掩码操作，缺失区域为 1，其他区域为 0。

**划分为子块**。我们将地图 按网格划分为规则的非重叠二维子块 ，然后我们将随机缺失的子块记为 ，剩余区域为 。

#### 2.2.2.2 STA 编码器

**线性嵌入**。我们的编码器通过可训练的线性投影 对未遮蔽的子块进行嵌入：

其中， 表示位置嵌入， 表示时间嵌入。

**分离的时空注意力模块**。首先，STA 编码器通过一系列分离的时空 Transformer 块处理嵌入向量 ，其中时间和空间注意力依次分离进行。对于每一层 ，我们计算查询、键、值向量：

其中， 表示层归一化， 表示多个注意力头的索引， 表示注意力头数，潜在维度为 。

分离时序注意力通过点积计算。对于每个块 ，我们首先计算 中的每个块以及在其他时间处于相同空间位置的所有块的时间注意力，如下所示：

其中， 表示 激活函数， 表示每个子块与时间维度上的所有同位置子块进行 次比较。然后，将这些来自所有注意力头的矢量拼接如下：

接着，作为新编码送入空间注意力模块。根据上式从 中重新获得 ，我们同时计算时间步 中每个子块与所有不同空间位置子块之间的空间注意力，具体如下：

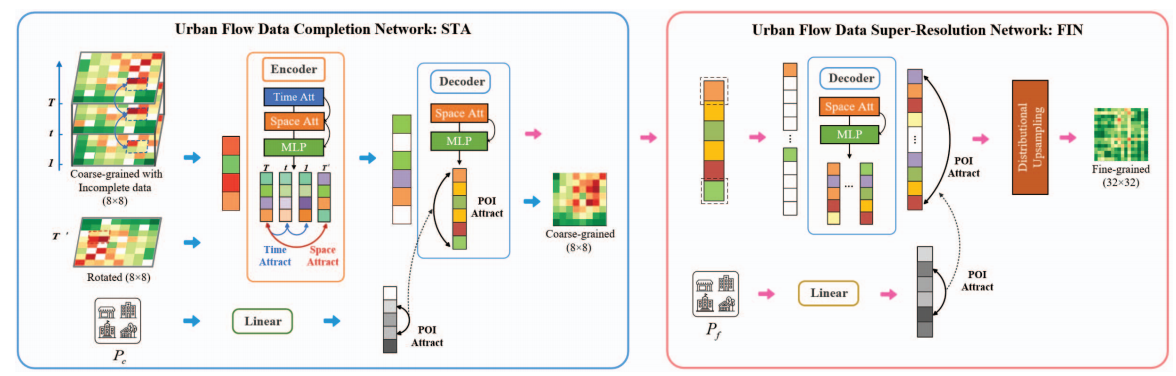


图 2 UrbanSTA 框架结构图。该方法由两个主要部分组成：时空注意力补全网络（STA）与细粒度推理网络（FIN）。STA 包含非对称的编码器与解码器，FIN 仅包含解码器。UrbanSTA 采用三阶段结构，将多任务问题分解为时空特征编码与解码步骤，以实现流量补全与超分辨率推理。

其中 表示每个子块在空间维度被比较 次。注意，分离注意力每个子块仅需 次比较，较其他自注意力计算方法更加高效[17]。

最终，所有 向量按照上述方式拼接，并通过多层感知机（MLP）进一步处理，结合残差连接获得 。

**时间与空间吸引**。我们在时间维度设计时间周期性损失 ，用于提取周期性信息。当前时刻 可以从历史数据中提取三个时间属性：接近性（closeness）、周期性（period）与趋势性（trend）。时间周期性损失定义为：

其中 表示周期性与趋势性特征图， 与 分别为周期性与趋势性时间片数。

为提升 Transformer 的局部特征捕捉能力，我们提出空间旋转损失 ，基于一定程度的旋转不变性。

将粗粒度流量图 随机旋转角度 得到 ，再分块并计算其空间注意力嵌入特征为 ，定义如下：

其中 表示旋转还原函数，用于将旋转后的特征恢复至原始位置。

#### 2.2.2.3 STA 解码器

**重建目标**。STA 解码器的输入是完整的粗粒度城市流量图子块集合 ，该集合包括编码后的未遮蔽子块 与掩码子块 ：

其中 表示较小维度的线性投影层， 为排序函数，用于将所有子块恢复至初始 的顺序， 为解码器的位置嵌入。

随后，向量 被输入至空间注意力模块以捕捉空间信息 。解码器特征经线性投影并重构为完整的粗粒度城市流量图 ：

其中 为 reshape 函数，将子块重构为二维图， 为线性投影矩阵。

**POI 吸引**。我们收集每个区域的兴趣点（POI）类别分布（如购物、医疗、教育和居民），记作粗粒度和细粒度的 POI 特征为 与 ，其中 表示兴趣点类别数[18]。

我们使用线性投影融合 类 POI，随后（如图 2 蓝色框底部）在粗粒度 POI 图中选取锚点区域 ，从候选区域 中依据距离与阈值选择 个正样本 与 个负样本 。

我们提出 POI 对比损失 ，旨在拉近正样本区域对应的流量特征 ，同时拉远负样本 ，定义如下：

其中 表示两个嵌入特征之间的相似度函数（例如内积）。

### 2.2.3 城市流量数据补全网络

**细粒度重构**。粗粒度特征 的关系需要被重构为细粒度特征关系，定义如下：

其中 表示学习得到的向量，代表缺失的子块。我们为每个粗粒度子块向量补充 个细粒度子块向量，从而形成细粒度关系映射图。共计 组 个子块向量作为输入。

**超分辨率推理**。随后，我们通过空间注意力模块进一步提取特征。粗粒度子块被丢弃，仅保留剩余 组细粒度子块，重组为二维图 ：

其中 表示按照区域位置将一维子块组合为二维图的重构函数， 表示线性投影。与此同时，我们选择 归一化策略，使细粒度流量之和等于其对应的粗粒度流量，定义如下：

其中 表示细粒度城市流量的分布概率。

### 2.2.4 最终目标函数

为了同时实现城市流量数据补全与超分辨率推理，我们提出了一个多任务学习网络，并采用两阶段的训练策略。

第一阶段为预训练阶段，针对数据补全网络，使用像素级均方误差（MSE）损失、时间周期损失、空间旋转损失与 POI 对比损失进行联合训练，定义如下：

其中 为超参数， 表示时间步 的粗粒度城市流量图真实值。

第二阶段训练中，我们在训练经过预训练的 STA 模块的基础上，引入 FIN 模块，加入像素级 MSE 损失与 POI 对比损失，最终实现最优总体UrbanSTA如下：：

其中 为超参数， 表示时间步 的细粒度城市流量图真实值。

## 2.3 实验与分析

### 2.3.1 实验设置

### 2.3.1.1 数据集

数据集的统计信息如表 I 所示。在实验中，我们将数据划分为互不重叠的训练集、验证集和测试集，比例为 2:1:1。

表 1 数据集统计信息

|  |  |  |
| --- | --- | --- |
| **数据集** | TaxiBJ | BikeNYC |
| 时间跨度 | 7/1/2013 - 10/31/2013 | 1/1/2019 - 3/31/2019 |
| 时间间隔 | 30分钟 | 1小时 |
| 粗粒度 | 8 × 8 | 8 × 8 |
| 细粒度 | 16 × 16 / 32 × 32 | 16 × 16 / 32 × 32 |
| 放大因子（M） | 2 / 4 | 2 / 4 |
| POI信息 | √ | √ |

### 2.3.1.2 对比方法

我们将 UrbanSTA 与六种基线方法进行了比较。其中，UrbanFM、UrbanPy、FODE 为细粒度流量推理方法；MT-CSR、STA-UP、STA-FD 为考虑缺失数据的细粒度流量推理方法。值得注意的是，STA-UP 和 STA-FD 分别是将 STA 模块与 UrbanPy 和 FODE 结合后的版本。

### 2.3.1.3 训练细节与超参数

子块大小设置为 1。STA 模型预训练时的学习率设为 ，批大小为 256；UrbanSTA 完整模型训练时的学习率为 ，批大小为 128。

### 2.3.2 与基线方法的对比分析

**性能对比**。不同方法的计算结果汇总如表 II 和表 III 所示。实验结果表明，UrbanSTA 在 TaxiBJ 数据集上的多数指标中取得了最佳性能。相比之下，UrbanFM、UrbanPy 和 FODE 忽略了数据缺失问题。而在 BikeNYC 数据集上，UrbanSTA 在多数情形中也获得了次优表现。由于 UrbanSTA 模型采用 Transformer 网络结构，在小规模数据集上的表现略逊于基于 CNN 的方法。

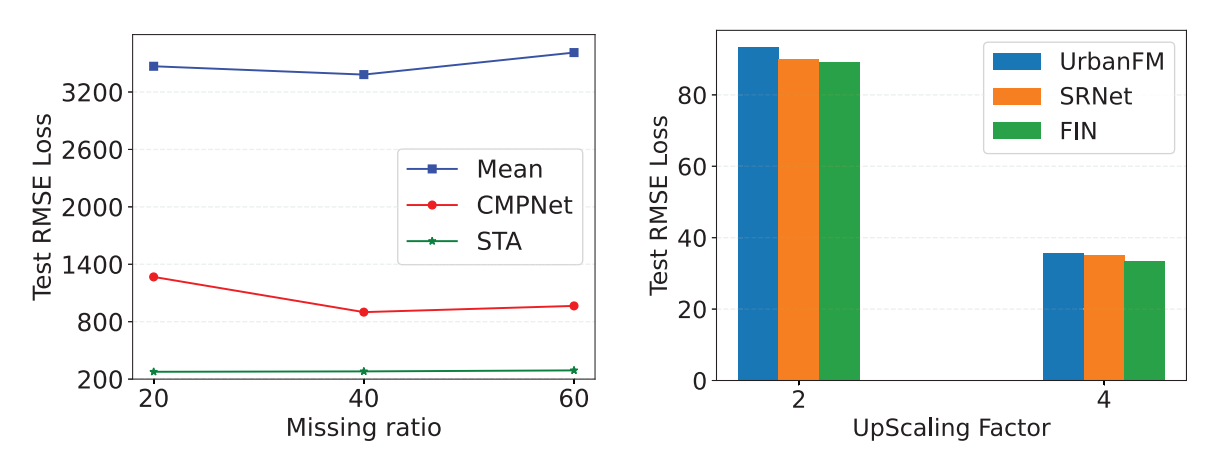


图 3 （a）STA 模块在任务中的有效性分析；（b）FIN 模块在任务中的有效性分析。

**模型在两项任务上的效果分析**。我们在 TaxiBJ 数据集上将 STA 模块与平均填补（Mean）和 CMPNet 进行了对比。如图 3(a) 所示，STA 模块明显优于两个基线方法。FIN 模块是一个不依赖数据补全的超分辨率模块。我们使用 TaxiBJ 数据集将 FIN 与 UrbanPy 及 SRNet 进行了比较，结果如图 3(b) 所示。结果表明，在不同上采样因子条件下，FIN 的表现始终略优于 UrbanPy 和 SRNet。

表 2 TaxiBJ 数据集上的对比实验结果。2 和 4 表示上采样因子，20%、40%、60% 表示数据缺失率。最佳结果加粗，次佳结果下划线标记。

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 模型 |  | | | UrbanFM | UrbanPy | FODE | MT-CSR | STA-UF | STA-FD | UrbanSTA |
| TaxiBJ | 2 | 20% | MAE | 73.25 | 118.87 | 74.39 | 75.81 | 67.62 | **62.99** | 63.03 |
| RMSE | 116.20 | 190.13 | 116.75 | 123.35 | 98.84 | 92.49 | **92.36** |
| 40% | MAE | 113.65 | 174.04 | 116.19 | 117.06 | 72.86 | 66.95 | **65.06** |
| RMSE | 174.22 | 255.15 | 176.04 | 186.86 | 105.26 | 98.32 | **96.01** |
| 60% | MAE | 132.48 | 190.76 | 133.62 | 134.32 | 73.64 | **68.89** | 69.91 |
| RMSE | 202.88 | 274.95 | 201.69 | 214.16 | 108.43 | **102.99** | 103.95 |
| 4 | 20% | MAE | 25.89 | 31.02 | 30.77 | 26.74 | 25.23 | 25.71 | **23.19** |
| RMSE | 40.67 | 48.26 | 46.54 | 44.71 | 38.69 | 40.03 | **37.11** |
| 40% | MAE | 35.53 | 43.50 | 39.59 | 36.34 | 25.84 | 24.84 | **24.36** |
| RMSE | 52.73 | 63.14 | 57.24 | 58.75 | 40.00 | 39.36 | **38.98** |
| 60% | MAE | 39.71 | 45.48 | 42.94 | 39.80 | 26.09 | 25.20 | **24.59** |
| RMSE | 60.65 | 66.24 | 62.45 | 65.20 | 40.87 | 40.49 | **39.95** |

表 3 BikeNYC 数据集上的对比实验结果。2 和 4 表示上采样因子，20%、40%、60% 表示数据缺失率。最佳结果加粗，次佳结果下划线标记。

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 模型 |  | | | UrbanFM | UrbanPy | FODE | MT-CSR | STA-UF | STA-FD | UrbanSTA |
| BikeNYC | 2 | 20% | MAE | 0.88 | 0.91 | 0.83 | 0.80 | 0.81 | 0.82 | **0.79** |
| RMSE | 2.17 | 2.24 | **1.82** | 1.89 | 2.05 | 1.84 | 1.87 |
| 40% | MAE | 1.07 | 1.20 | 1.09 | 1.02 | 1.08 | **0.87** | 1.44 |
| RMSE | 2.31 | 2.84 | 2.33 | 2.49 | 2.75 | **1.92** | 3.15 |
| 60% | MAE | 1.15 | 1.31 | 1.17 | 1.12 | 1.04 | 0.91 | **0.89** |
| RMSE | 2.48 | 3.23 | 2.49 | 2.58 | 2.24 | **2.00** | 2.09 |
| 4 | 20% | MAE | 0.27 | 0.28 | 0.27 | **0.26** | 0.33 | **0.26** | 0.36 |
| RMSE | **0.94** | 1.07 | **0.94** | 1.02 | 0.97 | 0.95 | 1.03 |
| 40% | MAE | 0.33 | 0.35 | 0.33 | 0.30 | 0.34 | **0.28** | 0.30 |
| RMSE | 1.04 | 1.34 | 1.04 | 1.16 | 1.00 | 0.99 | **0.98** |
| 60% | MAE | 0.35 | 0.35 | 0.35 | 0.33 | 0.35 | **0.32** | **0.32** |
| RMSE | 1.10 | 1.34 | 1.10 | 1.25 | **1.02** | 1.06 | **1.02** |

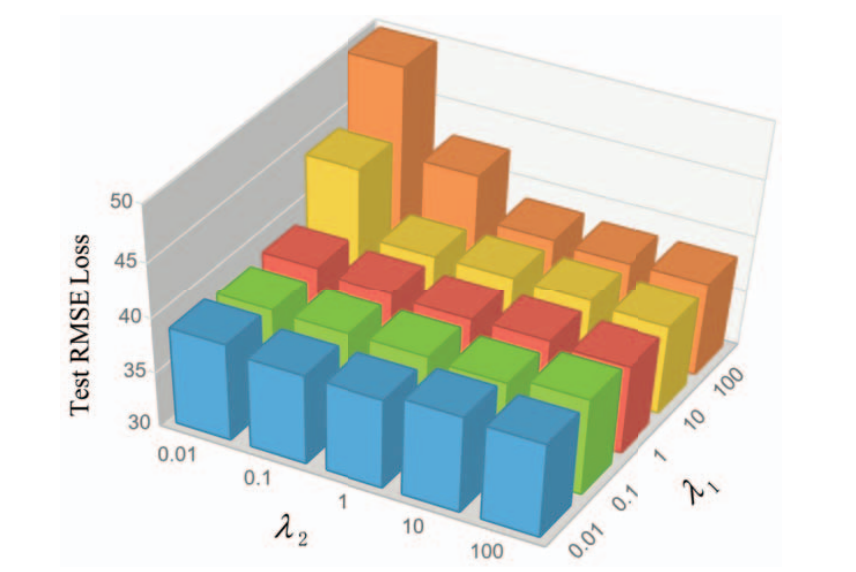


图 4 在 TaxiBJ 数据集（缺失率 20%，上采样因子 4）上的不同超参数设置下模型性能对比。

### 2.3.3 消融分析

我们在 TaxiBJ 数据集上对 UrbanSTA 及其若干变体进行了对比分析，计算结果如表 IV 所示。结果表明，STA 模块优于其七种变体，表明空间旋转、时间周期以及粗粒度 POI 对比损失在补全阶段均具有积极作用。

同时，我们还开展了对比实验以验证细粒度 POI 约束对 FIN 模块性能的影响。表 IV 总结了相关结果，添加细粒度 POI 约束显著提升了超分辨率推理任务的表现。

表 4 消融实验结果。在 TaxiBJ 数据集（缺失率 20%，上采样因子 4）上报告不同约束条件下模型性能。最佳结果加粗，次佳结果下划线标记。

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 空间 | |  | √ |  |  | √ | √ |  | √ |  | | √ | √ |
| 时间 | |  |  | √ |  | √ |  | √ | √ | √ | √ |
|  | |  |  |  | √ |  | √ | √ | √ | √ | √ |
|  | |  |  |  |  |  |  |  |  |  | √ |
| 数据补全 | MAE | 226.72 | 210.55 | 214.86 | 209.31 | 212.22 | 210.21 | 212.27 | 204.02 | 超分辨率 | MAE | 24.20 | 24.10 |
| RMSE | 307.90 | 289.40 | 290.67 | 287.84 | 288.76 | 287.21 | 288.07 | 274.80 | RMSE | 38.47 | 38.29 |

### 2.3.4 参数分析

我们分析了模型在不同参数设置下对 RMSE 指标的敏感性。图 4 总结了该实验结果，其中我们分别在 范围内调节最终目标函数中两个超参数 与 。实验结果显示，FIN 模块的作用大于 STA 模块。因此，在实际实验中，我们对 STA 模块进行轻微微调，增强对 FIN 模块的约束，最终设置为 ，，从而实现最佳性能。

# 4 结论（结论和建议）

本文提出了一种针对细粒度城市流量推理任务的神经网络模型 UrbanSTA，能够处理缺失数据情形。UrbanSTA 针对该任务中的三个关键挑战——粗粒度城市流量图的不完整性、粗细粒度流量数据间的复杂关联性以及城市流量图中空间属性的相互影响——提供了解决方案。我们在两个大规模城市流量数据集上进行了广泛评估，实验结果表明所提出的模型在性能上显著优于现有的先进方法。

# 参考文献

1. K. Su, J. Li, and H. Fu, “Smart city and the applications,” in 2011 International Conference on Electronics, Communications and Control (ICECC). IEEE, 2011, pp. 1028–1031.
2. H. R. Kwon, H. Cho, J. Kim, D. Lee, and S. K. Lee, “International case studies of smart cities: Namyangju, republic of korea,” 2016.
3. F. Zhou, L. Li, T. Zhong, G. Trajcevski, K. Zhang, and J. Wang, “Enhancing urban flow maps via neural odes,” in Proceedings of the Twenty-Ninth IJCAI, 2020.
4. Y. Gong, Z. Li, J. Zhang, W. Liu, B. Chen, and X. Dong, “A spatial missing value imputation method for multi-view urban statistical data,” in Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, 2021, pp. 1310–1316.
5. Y. Gong, X. Dong, J. Zhang, and M. Chen, “Latent evolution model for change point detection in time-varying networks,” Information Sciences, vol. 646, p. 119376, 2023.
6. J. Li, S. Wang, J. Zhang, H. Miao, J. Zhang, and P. Yu, “Fine-grained urban flow inference with incomplete data,” IEEE Transactions on Knowledge and Data Engineering, 2022.
7. Y. Gong, Z. Li, J. Zhang, W. Liu, Y. Yin, and Y. Zheng, “Missing value imputation for multi-view urban statistical data via spatial correlation learning,” IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 1, pp. 686–698, 2021.
8. Y. Gong, Z. Li, W. Liu, X. Lu, X. Liu, I. W. Tsang, and Y. Yin, “Missingness-pattern-adaptive learning with incomplete data,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.
9. P. Xie, T. Li, J. Liu, S. Du, X. Yang, and J. Zhang, “Urban flow prediction from spatiotemporal data using machine learning: A survey,” Information Fusion, vol. 59, pp. 1–12, 2020.
10. Y. Gong, Z. Li, J. Zhang, W. Liu, and Y. Zheng, “Online spatio-temporal crowd flow distribution prediction for complex metro system,” IEEE Transactions on knowledge and data engineering, vol. 34, no. 2, pp. 865–880, 2020.
11. L. Zhao, M. Gao, and Z. Wang, “St-gsp: Spatial-temporal global semantic representation learning for urban flow prediction,” in Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, 2022, pp. 1443–1451.
12. J. Zhang, Y. Zheng, D. Qi, R. Li, and X. Yi, “Dnn-based prediction model for spatio-temporal data,” in Proceedings of the 24th ACM SIGSPATIAL international conference on advances in geographic information systems, 2016, pp. 1–4.
13. J. Zhang, Y. Zheng, and D. Qi, “Deep spatio-temporal residual networks for citywide crowd flows prediction,” in Thirty-first AAAI conference on artificial intelligence, 2017.
14. H. Qu, Y. Gong, M. Chen, J. Zhang, Y. Zheng, and Y. Yin, “Forecasting fine-grained urban flows via spatio-temporal contrastive self-supervision,” IEEE Transactions on Knowledge and Data Engineering, 2022.
15. Y. Liang, K. Ouyang, L. Jing, S. Ruan, Y. Liu, J. Zhang, D. S. Rosenblum, and Y. Zheng, “Urbanfm: Inferring fine-grained urban flows,” in Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2019, pp. 3132–3142.
16. K. Ouyang, Y. Liang, Y. Liu, Z. Tong, S. Ruan, D. Rosenblum, and Y. Zheng, “Fine-grained urban flow inference,” IEEE transactions on knowledge and data engineering, 2020.
17. G. Bertasius, H. Wang, and L. Torresani, “Is space-time attention all you need for video understanding?” in Proceedings of the International Conference on Machine Learning (ICML), July 2021.
18. L. Wang, X. Geng, X. Ma, F. Liu, and Q. Yang, “Cross-city transfer learning for deep spatio-temporal prediction,” in International Joint Conference on Artificial Intelligence, 2019, p. 1893.