

DOI: 10.3969/j.issn.1671-0673. 2026. 03. 007

融合方位角信息的小视场角图像跨视角地理定位

张远东, 杨春芳, 孙思悦

(信息工程大学, 河南 郑州 450001)

摘要: 针对现有跨视角图像地理定位中视角变换易造成信息丢失、进而降低定位准确率的问题, 提出一种融合方位角信息的小视场角图像跨视角地理定位方法(FAECGeo)。首先, 设计融合方位角信息的特征调制模块, 利用图像上对应区域的方位角信息增强图像特征, 并预测方位角信息置信度图, 在增强特征与原始特征之间进行自适应调制融合, 以优化使用方位角信息与语义信息。然后, 设计基于方位角信息置信度加权的 LSE(LogSumExp)池化模块为累加的指数赋权, 以避免过于强调方位角信息对全局表征的干扰, 提升特征鲁棒性。实验结果表明, FAECGeo 在不同视场角条件下的定位准确率均优于现有先进方法, 如当视场角为 90° 时, 在 CVUSA 和 CVACT 数据集上的 R@1 最高分别提升 3.7 和 3.2 个百分点。

关键词: 跨视角图像地理定位; 小视场角图像; 方位角信息置信度; LogSumExp 池化

中图分类号: TP751

文献标识码: A

文章编号: 1671-0673(2026)03-0297-07

Cross-View Geo-Localization Method for Limited Field-of-View Images with Fused Azimuthal Information

ZHANG Yuandong, YANG Chunfang, SUN Siyue

(Information Engineering University, Zhengzhou 450001, China)

Abstract: To address the information loss caused by viewpoint changes in cross-view image geo-localization, which degrades localization accuracy, a cross-view geo-localization method for limited field-of-view images with fused azimuthal information (FAECGeo) is proposed. Firstly, a feature modulation module with fused azimuthal information is designed to enhance image feature maps from the corresponding image locations, while predicting an azimuthal information confidence map to adaptively fuse the enhanced features with the original ones, so as to optimize the utilization of azimuthal information and semantic information. Then, an LogSumExp (LSE) pooling module weighted by azimuthal information confidence is introduced to reweight the accumulated exponentials, reducing interference caused by overemphasized azimuthal information in global representations and improving feature robustness. Experimental results show that the localization accuracy of FAECGeo outperforms existing state-of-the-art methods under different field-of-view settings. When the field-of-view is 90° , the maximum R@1 scores on the CVUSA and CVACT datasets are improved by 3.7 and 3.2 percentage points, respectively.

Key words: cross-view image geo-localization; limited field-of-view images; azimuthal information confidence; LogSumExp pooling

跨视角图像地理定位(Cross-view Image Geo-localization)旨在通过匹配地面查询图像与带有位

置信息的卫星图像, 来确定地面图像的拍摄位置^[1-3]。该技术可被用于位置数据取证^[4]、虚假 GPS

收稿日期: 2026-02-18; 修回日期: 2026-03-18

基金项目: 国家自然科学基金(62472440); 河南省自然科学基金(252300421064)

作者简介: 张远东(1999—), 男, 硕士生, 主要研究方向为计算机视觉、图像检索和跨视角图像地理定位。E-mail: 19851897171@163.com

信息核验^[5]以及地理空间隐私保护^[6]等场景。

当前,跨视角图像地理定位的研究主要依赖深度学习网络进行特征提取^[7-9]。早期方法多采用卷积神经网络(Convolutional Neural Network, CNN)作为特征提取器,如Shi等^[8]使用孪生CNN提取图像特征。近年来,研究者们提出一系列基于Transformer的方法,如Zhu等^[9]提出的TransGeo利用全局注意力机制与位置编码捕捉跨视角特征对应关系。

上述研究主要关注地面全景图像定位。在实际应用中,地面图像多由手机、监控摄像头等设备获取,并不是全景图像形式,仅能覆盖局部区域。这使得小视场角图像跨视角地理定位成为一个更具挑战性的研究方向。

现有方法主要通过训练策略的改进和图像视角对齐来提升小视场角图像的跨视角定位准确率^[10-13]。如Shugaev等^[10]将卫星图像转换为极坐标表示并融合方向信息,提升在地面视场角受限且朝向未知条件下的跨视角特征对齐效果;Hu等^[14]结合双重加权软边界三元组损失与难负样本挖掘对特征进行约束,实现小视场角地面图像定位;Mi等^[15]引入多尺度数据增强,并提出单视角与跨视角结合的鲁棒对齐学习框架,以提升模型对视场角变化的适应能力;Cai等^[16]通过跨数据集扩充构造辅助训练集,并通过多视图对比学习和无人机图像监督,以增强有限视场角下的跨视角定位性能。这些措施一定程度上缓解了图像视场角范围受限与视角差异带来的匹配困难,尤其是通过图像变换的视角对齐方法增强了跨视角图像特征的对应关系。然而,图像变换会不可避免地造成部分信息丢失与扭曲形变,提取的图像特征可辨识性不足,容易产生误匹配,进而降低定位准确率。

针对上述挑战,本文提出了融合方位角信息的小视场角图像跨视角地理定位方法(Cross-View Geo-Localization with Fused Azimuthal Information, FAECGeo)。该方法将地面图像与卫星图像对应区域的方位角信息作为方向先验,注入图像特征图中,在无需额外视角变换与重采样的前提下实现特征增强,从而避免视角对齐带来的信息丢失与几何形变,优化使用方位角信息与语义信息。此外,本文设计了基于方位角信息置信度的加权LSE(LogSumExp)池化,使得到的全局描述更符合小视场角条件下的跨视角匹配需求,确保特征的鲁棒性。实验结果表明,FAECGeo方法在不同视场角条件下均显著提升了定位准确率。

1 提出的方法

针对现有方法中常用的图像变换不可避免地会导致部分信息丢失和几何形变、造成图像特征的辨识性下降的问题,本文提出了融合方位角信息的小视场角图像跨视角地理定位方法FAECGeo。该方法由4个结构相似的图像特征提取分支组成,输入分别为地面全景图像、小视场角地面图像、卫星图像和旋转卫星图像。

在每个分支中,首先,对输入的图像进行数据增强操作,以模拟视场角变化和方向扰动,并将增强得到的图像输入到特征提取骨干网络ConvNeXt,提取图像的原始特征图 F ;其次,利用本文设计的方位角信息增强块,将由方位角信息编码器生成的方位角信息特征向量注入特征图 F 中,得到包含方位角信息的增强特征图,并基于置信度预测头得到置信度图,利用置信度图对增强特征图与原始特征图进行自适应融合,生成方位角信息融合的特征图 F_{out} ;最后,利用设计的基于方位角信息置信度的加权LSE池化模块,对 F_{out} 进行置信度加权聚合,通过置信度权重抑制低置信度区域对全局表征的干扰,并采用加权LSE聚合突出高置信度区域内的判别性特征,从而得到用于跨视角地理定位的全局特征表示。

训练阶段采用四分支联合学习,同时输入地面全景图像、小视场角地面图像、卫星图像及其旋转卫星图像。并按图1所示在4个分支之间分别计算InfoNCE损失,包括地面全景图像与卫星图像、地面全景图像与小视场角地面图像、卫星图像与旋转卫星图像、小视场角地面图像与卫星图像4项损失。最终对各项损失加权求和进行优化。

其中,地面全景图像与卫星图像的对齐监督主要用于引导模型学习与小视场角图像覆盖区域相关的判别性特征表示。地面全景图像与小视场角图像的损失主要是利用地面全景图像对特征提取模型进行监督,使其能够根据小视场角地面图像中的局部信息推断出更完整内容的特征表示。卫星图像与旋转卫星图像之间的损失通过度量图像之间的特征差异,能够让模型从卫星图像中提取到对方向变化鲁棒的特征表示。测试阶段仅需输入小视场角地面图像和卫星图像,计算二者特征的余弦相似度,并据此检索确定待定位图像最可能的地理位置。FAECGeo方法具体框架如图1所示。

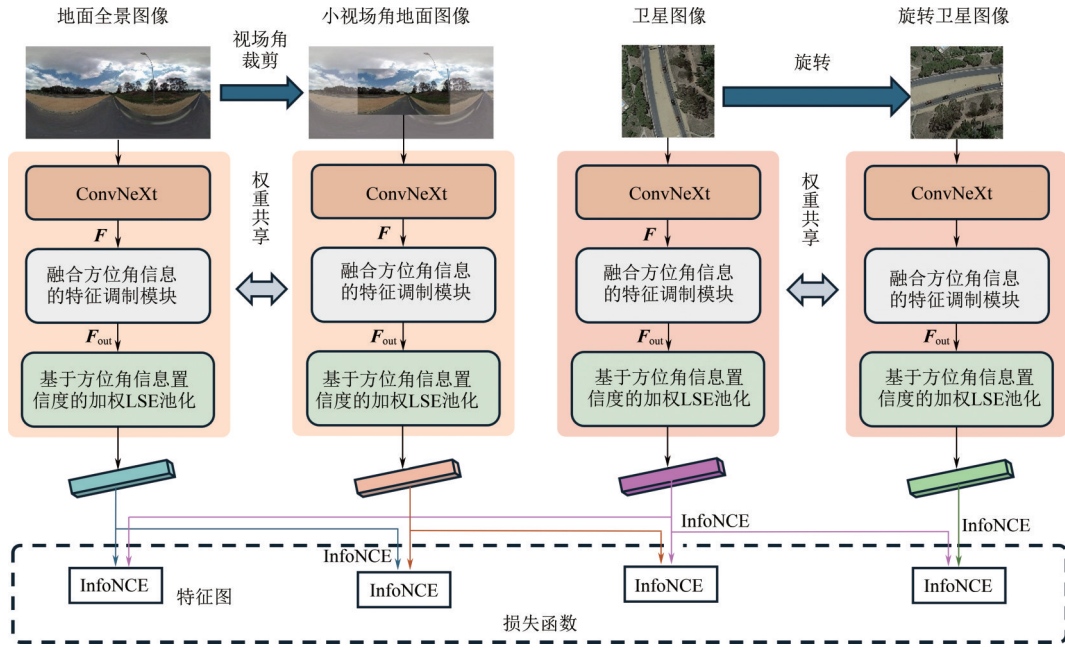


图1 FAECGeo方法框架

2 融合方位角信息的特征调制模块

为综合利用方位角信息与语义信息,本文设计了融合方位角信息的特征调制(Fused Azimuthal Information Feature Modulation, FAFM)模块,如图2所示。通过引入与视角无关的方位角信息,对图像特征图进行信息增强,并利用一个特征置信度图对方位角信息增强的特征图和原特征图进行调制,得到最终的图像特征图。

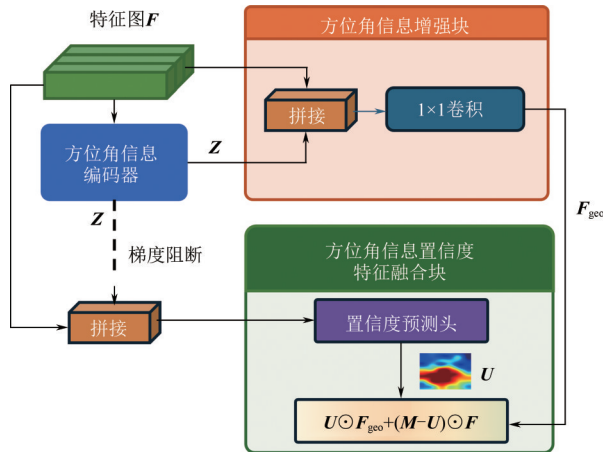


图2 融合方位角信息的特征调制模块

在FAFM模块中,首先利用一个方位角信息编码器,对输入图像对应区域的方位角进行编码,并将其用特征向量表示,组成图像的方位角信息图 Z ;接着,通过方位角信息增强块将方位角信息图 Z 与输入的图像特征图 F 进行拼接,并使用 1×1 卷积实现特征图融合与信息交互,得到注入方位角信

息的图像特征图 F_{geo} ;最后,通过置信度感知的特征调制块预测置信度图 U ,用于“方位角信息增强特征图 F_{geo} ”与“原始特征图 F ”间的调制,从而使最终的匹配能够综合利用方位角信息与语义信息。

2.1 方位角信息编码器

输入图像的位置坐标归一化表示为 (x, y) ,其中 x 与 y 分别由像素的水平/垂直坐标按图像宽度 W 与高度 H 线性映射到 $[-1, 1]$ 。考虑到地面图像与卫星图像在视角上差异显著,而方位角 θ 在两种视角下具有一致的方向语义。因此,本文使用方位角信息编码器将图像特征图上的归一化坐标转换为对应的方位角信息,用于后续的特征增强。

当输入图像为卫星图像时,卫星图像中每个位置的方位角可表示为

$$\theta = \text{atan2}(y, x) \quad (1)$$

式中, $\text{atan2}(\cdot, \cdot)$ 表示反正切函数,用于计算方位角 θ 。

当输入图像为地面图像时,考虑到地面图像视场角对方位角角度范围的影响,方位角由归一化坐标 x 线性映射得到:

$$\theta = \frac{1}{2} R_{\text{FoV}} x \quad (2)$$

式中, $R_{\text{FoV}} \in [0, 2\pi]$ 为地面图像的视场角大小。

方位角信息编码器通过傅里叶频率展开,将方位角 θ 编码为正余弦张量 $P(\theta)$,其形式为

$$P(\theta) = [\sin(2^0\theta), \cos(2^0\theta), \dots, \sin(2^K\theta), \cos(2^K\theta)] \quad (3)$$

式中, K 为频率个数。随后,通过一个轻量卷积投影将 $P(\theta)$ 映射为方位角信息特征图 Z :

$$\mathbf{Z} = f_{\text{proj}}(\mathbf{P}(\theta)), \mathbf{Z} \in \mathbb{R}^{B \times C_\theta \times H \times W} \quad (4)$$

式中: $f_{\text{proj}}(\cdot)$ 表示用 1×1 卷积进行逐位置线性映射; B 为批次大小; H, W 为特征图尺寸; C_θ 为输出通道数。

2.2 方位角信息增强块

方位角信息增强块主要将方位角信息编码器生成的方位角信息特征 $\mathbf{Z}$ 注入特征图中, 使特征图兼具语义信息与方位角信息, 更有利于跨视角场景下的特征对齐与匹配。设主干网络提取的特征图为 $\mathbf{F} \in \mathbb{R}^{B \times C \times H \times W}$ 。在通道维度将特征图 $\mathbf{F}$ 与方位角信息特征 $\mathbf{Z}$ 进行拼接, 并采用 1×1 卷积进行跨通道特征融合, 得到注入方位角信息后的增强特征 $\mathbf{F}_{\text{geo}}$:

$$\mathbf{F}_{\text{geo}} = f_{\text{fuse}}(\text{Concat}(\mathbf{F}, \mathbf{Z})) \quad (5)$$

式中: $\text{Concat}(\cdot)$ 表示对特征进行通道维度的拼接; $f_{\text{fuse}}(\cdot)$ 为 1×1 卷积融合方法。

上述处理方式无需引入额外的图像变换或重采样操作, 有效降低了因视角差异带来的特征错位风险。

2.3 方位角信息置信度特征融合块

在实际定位中, 由于图像中方位角信息与内容信息所起作用不尽相同, 为尽可能优化利用方位角信息与语义信息, FAFM 模块引入方位角信息置信度特征融合块。利用不同区域方位角信息增强后特征图的置信度, 对增强特征图和原始特征图进行调制, 以融合得到更有效的特征图。

首先, 通过将方位角信息特征图 $\mathbf{Z}$ 与经过梯度阻断的原始图像特征图 $\mathbf{F}$ 在通道维度拼接, 经一个 1×1 卷积层输出方位角信息置信度图 $\mathbf{U}$:

$$\mathbf{U} = \sigma(g(\text{Concat}(\mathbf{Z}, \text{sg}(\mathbf{F})))) \quad (6)$$

式中: $\mathbf{U} \in \mathbb{R}^{B \times 1 \times H \times W}$, 其中每个元素 $U_{b,1,h,w}$ 的取值范围为 $[0, 1]$; $g(\cdot)$ 表示置信度预测头; $\sigma(\cdot)$ 为 Sigmoid 函数; $\text{sg}(\cdot)$ 为禁止梯度操作。该预测头生成图像对应的方位角信息置信度值, 并使用 Sigmoid 函数将置信度值映射到 $[0, 1]$ 范围内。为使网络稳定学习置信度图, 在置信度图预测过程中, 对输入的原始特征 $\mathbf{F}$ 进行梯度阻断, 以提高置信度学习的稳定性, 作为一个固定的调节权重用于特征融合。

随后, 通过 $\mathbf{U}$ 逐区域加权调制增强特征图 $\mathbf{F}_{\text{geo}}$ 与原始图像特征图 $\mathbf{F}$, 得到输出特征 $\mathbf{F}_{\text{out}}$:

$$\mathbf{F}_{\text{out}} = \mathbf{U} \odot \mathbf{F}_{\text{geo}} + (\mathbf{M} - \mathbf{U}) \odot \mathbf{F} \quad (7)$$

式中: $\odot$ 表示逐元素相乘; $\mathbf{M} \in \mathbb{R}^{B \times 1 \times H \times W}$ 表示与 $\mathbf{U}$ 同尺寸的元素全为 1 的张量。该形式使得在信息置信度较高的位置, 增强特征得到更大权重, 以强化方位角信息对跨视角匹配的贡献; 在置信度较

低的位置, 模型更倾向于保留原始语义表征, 以降低方位角信息不可靠时带来的干扰, 从而提升跨视角图像特征匹配的鲁棒性。

3 基于方位角信息置信度的加权 LSE 池化模块

在现有的池化方法中, LSE 池化提供了一种平滑过渡的方式, 兼具平均池化的稳定性与最大池化的判别特性, 从而更好地平衡特征选择与聚合的稳定性。在此基础上, 进一步引入第 2.3 节得到的方位角信息置信度, 使聚合过程能够优化使用不同区域的方位角信息和语义信息, 减少过于强调方位角信息对全局表征的干扰。

如图 3 所示, 设 FAFM 输出特征图 $\mathbf{F}_{\text{out}}$ 中元素为 $f_{c,h,w}$, 置信度 $\mathbf{U}$ 作为权重参与池化计算, 其元素为 $U_{h,w}$, 对每个通道 c , 基于方位角信息置信度的加权 LSE 池化模块(CG-WLSE)首先在空间维度进行置信度加权的归一化聚合。

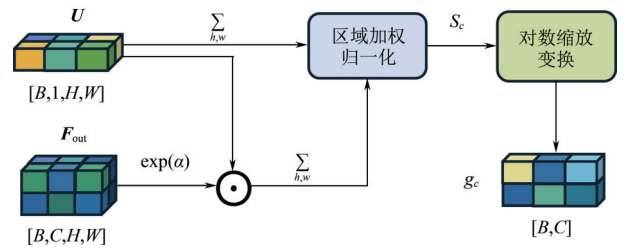


图3 基于方位角信息置信度的加权 LSE 池化模块

$$S_c = \sum_{h=1}^H \sum_{w=1}^W U_{h,w} \exp(\alpha f_{c,h,w}) \quad (8)$$

式中, α 为温度参数, 用于调节 LSE 池化在“平均”与“最大”之间的平滑程度。随后, 对聚合结果进行归一化并施加对数变换, 得到通道级全局响应。对聚合结果施加对数缩放得到全局通道响应:

$$g_c = \frac{1}{\alpha} \log_2 \left(S_c / \sum_{h=1}^H \sum_{w=1}^W U_{h,w} \right) \quad (9)$$

全局描述特征向量记为 $\mathbf{g} = [g_1, \dots, g_c] \in \mathbb{R}^{B \times C}$ 。

该聚合过程将置信度 $\mathbf{U}$ 作为权重, 抑制低置信度区域中方位角信息对全局表征的影响, 并利用 LSE 池化的平滑特性突出高置信度区域的判别性特征, 使得到的全局特征更适用于跨视角图像特征的匹配。

4 实验结果与分析

本节对给出的定位方法 FAECGeo 进行实验测试, 并对涉及融合方位角信息的特征调制模块和

基于方位角信息置信度的加权LSE池化模块进行消融实验,以分析它们对方法性能的影响。

4.1 实验数据集与评估指标

实验采用跨视角图像地理定位领域常用的两个公开数据集CVUSA和CVACT。CVUSA数据集覆盖美国多个地区,包含了35 532对训练样本与8 884对测试样本;卫星图像分辨率为 750×750 ,地面全景图像分辨率为 $224 \times 1\,232$ 。CVACT数据集主要覆盖澳大利亚堪培拉地区,数据集规模与CVUSA一致,卫星图像分辨率为 $1\,200 \times 1\,200$,地面全景图像分辨率为 $832 \times 1\,664$ 。

实验采用跨视角地理定位领域广泛使用的Recall@K(Top-K召回率)作为评估指标。该指标定义为:给定一个待查询的地面图像,检索与其特征相似度最高的前K张卫星图像。若其中包含与查询图像地理位置完全匹配的正样本卫星图像,则视为检索成功。本文选取常用的R@1、R@5、R@10及R@1%(在前1%检索结果中的命中率)4个指标,以衡量模型的跨视角定位性能。

4.2 实验环境及参数设置

本文方法基于PyTorch实现,并部署在48 GB NVIDIA RTX A6000 GPU上进行训练与测试。主干网络采用ConvNeXt-B,并使用ImageNet预训练权重进行初始化。优化器采用AdamW,初始学习率设为0.000 1,并使用余弦学习率调度器进行衰减。温度参数 α 为可学习标量,采用对数参数初始化,初始化值为1.0。为适配特征提取网络的输入尺寸,将CVUSA、CVACT数据集中的地面图像和卫星图像大小分别调整为 139×768 和 384×384 。每组实验训练60个轮次,批次大小设为32。旋转卫星图像由卫星图像随机顺时针旋转 90° 、 180° 、 270° 得到。

为贴近真实拍摄场景,本文选取 70° (手机相机)与 90° (广角相机)两类常见相机拍摄视场角大小^[2],并据此生成小视场角地面图像。现有研究中,常见做法是从地面全景图像中按目标视场角直接进行窗口裁剪,然后在此基础上开展训练与测试。然而直接裁剪得到的小视场角图像难以充分反映真实相机成像过程中的几何变化。相比之下,采用透视投影从全景图像中生成小视场角图像,在几何形变上更符合真实相机成像过程,如图4所示。因此,实验将在采用透视投影方式生成的小视场角图像上进行训练和测试。

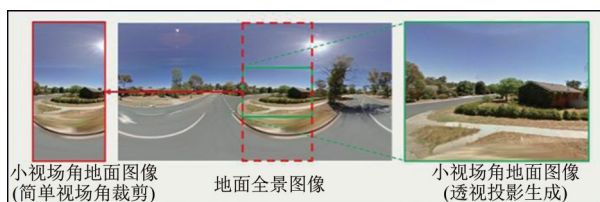


图4 两种小视场角地面图像的生成方式对比

4.3 实验结果分析

4.3.1 对比实验分析

为测试本文方法对小视场角地面图像的跨视角地理定位性能,在上述两个公开数据集上,将本文方法与现有先进方法GeoDTR^[17]、TransGeo^[9]、SAIG-D^[18]、Sample4Geo^[19]、ConGeo^[15]进行对比实验。表1和表2分别给出了不同方法在CVUSA和CVACT两个数据集上,采用透视投影生成的小视场角地面图像时, 90° 与 70° 视场角设置下的定位准确率对比结果。

表1 各方法在CVUSA数据集上的定位准确率 单位:%

方法	90°				70°			
	R@1	R@5	R@10	R@1%	R@1	R@5	R@10	R@1%
GeoDTR	9.0	20.1	27.8	61.6	5.8	12.8	18.1	46.7
TransGeo	17.8	34.9	44.4	74.3	10.7	24.1	31.2	60.1
SAIG-D	18.8	36.6	46.3	75.9	11.2	24.3	31.7	60.6
Sample4Geo	21.6	44.2	54.3	83.7	7.6	20.5	29.6	64.6
ConGeo	31.4	52.9	62.1	84.3	12.9	27.4	35.9	66.8
本文方法	35.1	54.6	63.6	85.4	15.4	29.2	36.7	67.4

表2 各方法在CVACT数据集上的定位准确率 单位:%

方法	90°				70°			
	R@1	R@5	R@10	R@1%	R@1	R@5	R@10	R@1%
GeoDTR	5.8	15.6	22.7	56.3	3.9	11.9	17.4	48.1
TransGeo	8.0	21.6	29.9	64.1	4.3	13.2	18.9	51.5
SAIG-D	9.7	24.0	32.9	67.2	6.5	18.1	25.4	58.4
Sample4Geo	30.9	57.1	66.7	89.4	16.3	37.0	47.4	79.7
ConGeo	50.6	71.9	78.1	90.8	32.7	55.1	63.7	85.1
本文方法	53.8	73.0	78.9	91.2	36.2	59.7	66.5	85.8

由表1和表2可得,随着视场角的缩小,地面图像可见内容逐步减少,跨视角图像匹配难度显著增加,各方法的定位性能均出现明显下降。然而,本文方法在两个数据集上均取得了优于现有先进方法的定位准确率,特别是在关键评估指标R@1上提升较为显著。

在CVUSA数据集上, 90° 与 70° 两种设置下,本文方法的R@1分别达到35.1%与15.4%,相较于ConGeo分别提升3.7与2.5个百分点;在CVACT数据集上,本文方法的R@1分别达到53.8%与

36.2%,相较于 ConGeo 分别提升 3.2 与 3.5 个百分点。这表明,在更符合真实相机成像的小视场角图像条件下,本文方法仍能稳定提升跨视角定位准确率。

图 5 和图 6 分别给出视场角大小为 90° 和 70° 时,本文方法与 ConGeo 方法在 CVACT 和 CVUSA 数据集上的 Top-3 定位结果对比,其中红色边框表示卫星图像正样本。从可视化结果来看,本文方法在所有示例中均实现了 Top-3 准确检索,而 ConGeo 方法在多个示例的 Top-3 检索结果中将正确样本排在靠后位置。



图 5 视场角大小为 90° 时,本文方法与 ConGeo 方法在 CVACT 数据集上的 Top-3 定位结果

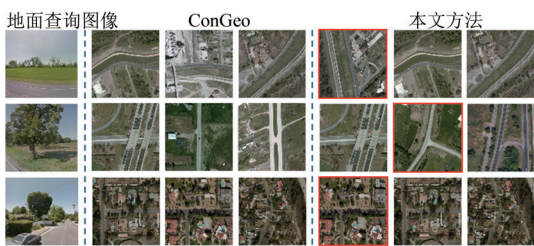


图 6 视场角大小为 70° 时,本文方法与 ConGeo 方法在 CVUSA 数据集上的 Top-3 定位结果

4.3.2 消融实验

消融实验均在采用透视投影生成的小视场角地面图像条件下进行,以进一步分析融合方位角信息的特征调制模块和基于方位角信息置信度的加权 LSE 池化模块对定位性能的影响。

1)融合方位角信息的特征调制模块消融实验。由于加权 LSE 池化模块的权重分配高度依赖于特征调制模块 FAFM 生成的置信度图 U ,因此,在进行特征调制模块消融实验时,同时去除加权 LSE 池化模块。表 3 给出了在透视投影生成的小视场角地面图像、视场角为 70° 条件下,融合方位角信息的特征调制模块在 CVUSA 和 CVACT 数据集上的消融实验结果。

实验结果表明,引入特征调制模块后,CVUSA 上的 R@1 由 12.9% 提升至 14.6%,提高了 1.7 个百分点;CVACT 上的 R@1 由 32.7% 提升至 35.2%,提高了 2.5 个百分点。这表明,融合方位角信息的特

征调制模块能够有效减小跨视角匹配中的方向偏差,提升模型的定位精度。

表 3 特征调制模块的消融实验结果 单位:%

方法	CVUSA		CVACT	
	R@1	R@1%	R@1	R@1%
BaseLine	12.9	66.8	32.7	85.1
BaseLine+FAFM	14.6	67.1	35.2	85.3

2)基于方位角信息置信度的加权 LSE 池化模块消融实验。为评估本文提出的基于方位角信息置信度的加权 LSE 池化模块(CG-WLSE)对模型定位性能的贡献,在透视投影生成的小视场角地面图像、视场角为 90° 的条件下,分别使用随机池化、最大池化、平均池化、LSE 池化以及 CG-WLSE 池化,在 CVUSA 和 CVACT 数据集上进行测试。

表 4 给出了不同池化方法的消融实验结果。可以看出,相较于随机池化、最大池化和平均池化,使用 LSE 池化和 CG-WLSE 池化后定位准确率得到了整体提升。由于 CG-WLSE 池化引入方位角信息置信度进行加权调节,能够根据不同区域的重要性自适应分配特征聚合权重,因此其定位性能最佳。

表 4 不同池化方法的消融实验结果 单位:%

方法	CVUSA		CVACT	
	R@1	R@1%	R@1	R@1%
随机池化	33.0	84.5	50.6	77.1
最大池化	34.9	85.1	52.3	78.3
平均池化	33.6	84.9	51.9	78.2
LSE 池化	34.2	85.2	53.1	78.6
CG-WLSE 池化	35.1	85.4	53.8	78.9

5 结束语

本文提出了一种融合方位角信息的小视场角图像跨视角地理定位方法 FAECGeo。通过设计融合方位角信息的特征调制模块,实现方位角信息与语义信息的有效融合,并使用设计的基于方位角信息置信度的加权 LSE 池化模块,兼顾平均池化的稳定性与最大池化的显著性选择能力,增强全局描述在复杂干扰与有限视场角条件下的鲁棒性。实验结果表明,所提方法在多种视场角条件下整体优于现有先进方法。

参考文献:

[1] LIN T Y, BELONGIE S, HAYS J. Cross-view image geo-

- localization [C] // Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2013: 891-898.
- [2] 周博文, 李阳, 马鑫骥, 等. 深度学习的跨视角地理定位方法综述[J]. 中国图象图形学报, 2024, 29(12): 3543-3563.
- [3] 赵璐颖, 周杨, 胡校飞, 等. 跨视角图像地理定位研究综述[J]. 地球信息科学学报, 2025, 27(10): 2293-2315.
- [4] GEPPERT M, LARSSON V, SPECIALE P, et al. Privacy preserving localization and mapping from uncalibrated cameras [C] // Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2021: 1809-1819.
- [5] HU S X, FENG M D, NGUYEN R M H, et al. CVM-net: cross-view matching network for image-based ground-to-aerial geo-localization [C] // Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2018: 7258-7267.
- [6] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision [C] // Proceedings of the 38th International Conference on Machine Learning. New York, USA: PMLR, 2021: 8748-8763.
- [7] 王彤, 纪文宇, 濮毅婷, 等. 物理感知增强的雾天跨视角地理定位方法[J/OL]. 计算机科学, [2026-06-25]. <https://link.cnki.net/urlid/50.1075.TP.20251219.1123.006>.
- [8] SHI Y J, YU X, CAMPBELL D, et al. Where am I looking at? joint location and orientation estimation by cross-view matching [C] // Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2020: 4063-4071.
- [9] ZHU S J, SHAH M, CHEN C. TransGeo: transformer is all you need for cross-view image geo-localization [C] // Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2022: 1152-1161.
- [10] SHUGAEV M, SEMENOV I, ASHLEY K, et al. ArcGeo: localizing limited field-of-view images using cross-view matching [C] // Proceedings of the 2024 IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway, USA: IEEE, 2024: 208-217.
- [11] WANG W, CUI J K, CAO X W, et al. Vaelnet: view augmentation contrastive learning for limited FOV cross-view geo-localization [C] // Proceedings of the 2025 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, USA: IEEE, 2025: 2177-2181.
- [12] RODRIGUES R, TANI M. Global assists local: effective aerial representations for field of view constrained image geo-localization [C] // Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway, USA: IEEE, 2022: 2694-2702.
- [13] LI C R, YAN C, XIANG X J, et al. AMPLE: automatic progressive learning for orientation unknown ground-to-aerial geo-localization [J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: No.5800115.
- [14] HU B N, LI H W, BU S H, et al. OriLoc: unlimited-FoV and orientation-free cross-view geolocalization [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 15508-15522.
- [15] MI L, XU C, CASTILLO-NAVARRO J, et al. ConGeo: robust cross-view geo-localization across ground view variations [C] // Proceedings of the 18th European Conference on Computer Vision. Cham, Switzerland: Springer, 2024: 214-230.
- [16] CAI K Y, LI X, TIAN F H, et al. GeoMatch: multi-view contrastive learning for limited CVGL with semantic uncertainty [C] // Proceedings of the 3rd International Workshop on UAVs in Multimedia: Capturing the World from a New Perspective. New York, USA: ACM, 2025: 10-14.
- [17] ZHANG X H, LI X Y, SULTANI W, et al. Cross-view geo-localization via learning disentangled geometric layout correspondence [C] // Proceedings of the 37th AAAI Conference on Artificial Intelligence. Palo Alto, USA: AAAI Press, 2023, 37(3): 3480-3488.
- [18] ZHU Y Y, YANG H J, LU Y X, et al. Simple, effective and general: a new backbone for cross-view image geo-localization [DB/OL]. (2023-02-03) [2026-02-09]. <https://doi.org/10.48550/arXiv.2302.01572>.
- [19] DEUSER F, HABEL K, OSWALD N. Sample4Geo: hard negative sampling for cross-view geo-localisation [C] // Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision. Piscataway, USA: IEEE, 2023: 16801-16810.

(编辑: 邵 华)

(上接第296页)

- [13] 肖秀丽, 唱亮. 测向交会定位的定位精度分析[J]. 中国无线电, 2011(4): 51-53; 66.
- [14] 杨思明, 单征, 丁煜, 等. 深度强化学习研究综述[J]. 计算机工程, 2021, 47(12): 19-29.
- [15] 祝亢, 黄珍, 王绪明. 基于深度强化学习的智能船舶航迹跟踪控制[J]. 中国舰船研究, 2021, 16(1): 105-113.
- [16] 万里鹏, 兰旭光, 张翰博, 等. 深度强化学习理论及其应用综述[J]. 模式识别与人工智能, 2019, 32(1): 67-81.

(编辑: 邵 华)