VIGOR：超越一对一检索的交叉视图图像地理定位

Sijie Zhu

摘要

摘要横向图像地理定位旨在确定通过匹配街景查询图像的位置 GPS 标记的鸟瞰参考图像。最近的关于图像检索的研究已经取得了惊人的高检索精度在城市规模的数据集上。然而，这些结果依赖于假设存在精确地以任何查询图像的位置为中心的参考图像，这不适用于实际场景。在本文中，我们重新定义这个问题有一个更现实的假设，即图像可以是感兴趣区域中的任意图像，并且在查询出现之前捕获参考图像。这假设打破了现有数据集的一对一检索设置，因为查询和参考图像不是并且可能存在覆盖一个查询位置的多个参考图像。桥接这种现实设置与现有数据集之间的差距，我们提出了一个新的大规模基准 VIGOR，用于一对一检索之外的交叉视图图像地理定位。我们对现有的最先进的方法进行了基准测试，并提出了一个新的端到端框架来本地化查询以粗到细的方式。除了图像级检索精度外，我们还评估了定位精度根据使用原始 GPS 的实际距离（米）数据在不同的应用场景下进行了大量实验，以验证提出的方法。结果表明，在这种现实环境中的横向地理定位仍然具有挑战性，这将促进这一方向的新研究。

关键词：图像检索；地理位置查询；一对多检索

1 引言

基于图像的地理定位的目标是通过找到最 GPS 标记的参考数据库中的类似图像。这样的事实证明，这些技术对精确定位非常有用带有嘈杂的 GPS 信号[1-2]和拥挤的导航城市[3-4]。最近，人们的兴趣激增在横截面地理定位[5-10]中使用 GPS 标记的鸟瞰图像作为街景查询的参考。但是，性能可能会受到查询和引用之间的视图或外观差距较大图像。最近的工作[7-9]表明，交叉视图图像匹配的性能可以显著提高通过特征聚合和样本挖掘策略改进。当街道视图（或地面视图）的方向图像可用（由基于电话的指南针提供），最先进的方法可以实现 80% 以上的顶级检索精度[8]，这表明了在真实世界环境中进行精确地理定位的可能性。然而，现有数据集[5,11-12]只是假设每个查询地面视图图像具有一个相应的参考空中视图图像，其中心在该位置精确对齐查询图像的。我们认为这对于现实世界的应用程序来说是不实际的，因为查询图像可能出现在感兴趣区域和参考中的任意位置应该在查询出现之前捕获图像。在里面在这种情况下，完全对齐的一对一对应关系是不保证。鉴于这个问题的新颖性，我们建议一个新的基准（VIGOR），用于在更现实的环境中评估交叉视图地理定位。简单地说，给定一个区域感兴趣（AOI），参考航空图像密集采样以实现 AOI 和在任意位置捕获街景查询。超越一对一：以前的研究主要集中在基于现有数据集的一对一对应将完全对齐的图像对视为默认值。然而 VIGOR 使我们能够探索参考样本的效果，这些样本不是以查询位置为中心覆盖查询区域。因此，可能存在多个部分覆盖相同查询位置的参考图像，打破了一对一的对应关系。在我们的地理定位方法中，我们设计了一种新的混合损耗在训练期间多个参考图像的优点。超越检索：图像检索只能提供图像级定位。由于中心对齐不是在检索之后，我们进一步使用图像内校准来预测检索到的图像内的查询位置。因此所提出的联合检索和校准框架提供了从粗到细的定位。整个管道是端到端，并且推断与偏移预测一样快与检索任务共享特征描述符。此外，我们的数据集还附有原始 GPS 数据。因此更直接的性能评估，可以

在我们的数据集上实现。我们的主要贡献概括如下：• 我们为交叉视图问题引入了一个新的数据集图像地理定位。这个数据集第一次，允许人们在更现实的情况下研究这个问题和实用的设置，并提供了桥接当前研究与实际应用之间的差距。• 我们提出了一种新的联合检索和校准框架，用于从粗到细的精确地理定位这是过去没有探索过的。• 我们开发了一种新的混合损失，以在训练期间从多个参考图像中学习在各种实验环境中有效。• 我们还验证了拟议交叉视图的潜力现实应用中的地理定位框架场景（辅助导航）。

2 VIGOR 数据集

问题描述：给定感兴趣的领域（AOI），我们的目标是在通过将其与航空参考图像进行匹配。为了确保任何可能的查询被至少一个参考图像覆盖，参考航空图像必须提供 AOI 的无缝覆盖。如图 1（a）所示，粗采样参考图像（黑色方框）无法提供 AOI 的全部覆盖范围查询位置（红星）可能位于参考样品。即使查询位置（黄色星形）位于参考航空图像的边缘，该参考图像仅与其中心位于查询位置，但可能不是提供足够的信息以区别于其他负参考图像。这些查询可以通过添加额外的重叠样本（绿色框）。像如图 1（b）所示，如果查询位置（红星）位于 ×L 航空图像的中心区域（黑色虚线框），查询和参考图像定义为正样品之间的相互作用。外部其他查询（蓝色星号）中心区域被定义为半阳性样本。到保证任意查询都有一个正引用图像，如图 1（c）所示，我们建议使用沿纬度和经度方向重叠 50%。

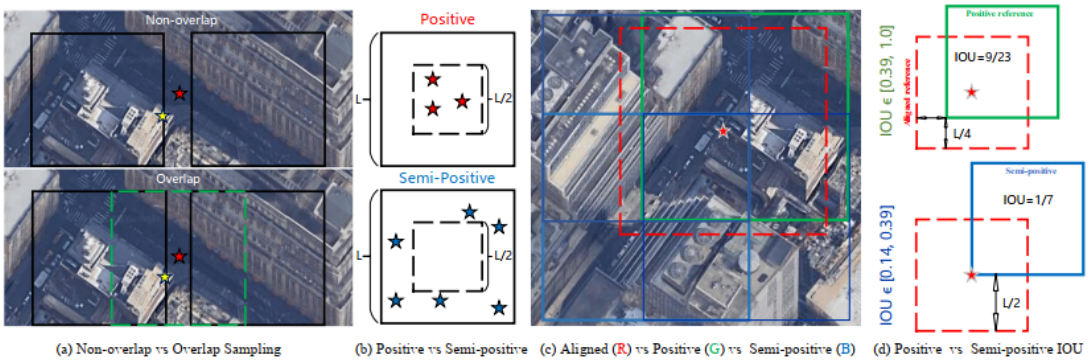

图 1: 建议数据集的采样策略（星号表示查询位置）

通过这样做，任何任意查询 AOI 中的位置（红星）由四幅参考图像（大小 L×L）覆盖。绿色框表示正参考和其他三个半正参考表示为蓝色框。正参考被认为是地面真实值，因为它具有距离查询并包含与查询共享最多的对象形象红色框表示完全对齐的航空图像。基于正和半正的定义如图 1（b）所示，我们可以很容易地看到参考图像有一个 IOU（在并集上相交）大于 0:39，与完全对齐的参考（参见图 1（d））。典型阳性样本的 IOU（相对于中心的偏移等于（±L/8；±L/8））为 0:62。半正样本和校准参考之间的差值为 [1/7≈0:14；9/23≈0:39]。

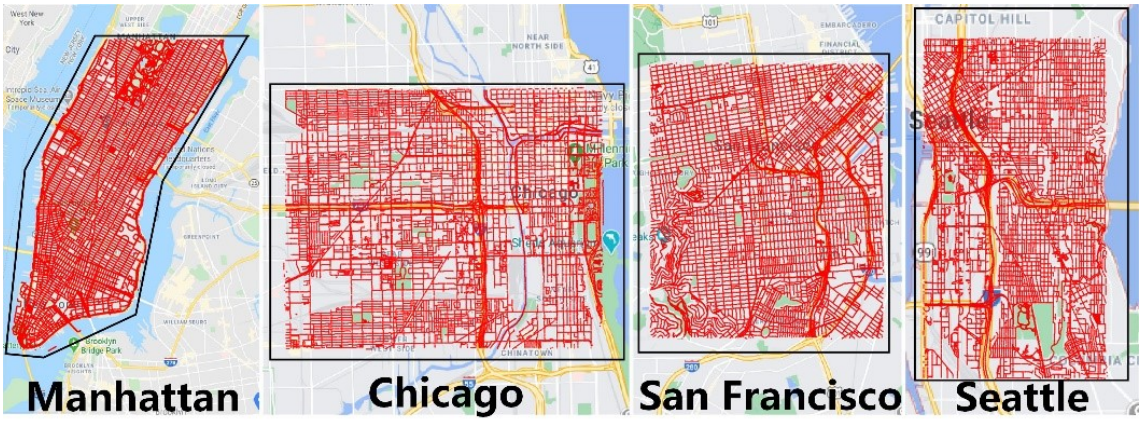

图 2: 四个城市的航空图像覆盖率（黑色多边形）全景图的分布（红点）

数据收集：如图 2所示，我们收集了 90；618 覆盖四个城市中心区域的航空图像，即。纽约市（曼哈顿）、旧金山、芝加哥和西雅图，作为使用谷歌地图静态 API 的 AOI[13]。然后 238；收集了 696 张街景全景图像在缩放级别为 2 时使用 Google 街景静态 API[14]在大多数街道上。全景的所有 GPS 位置图像在我们的数据集中是唯一的样本之间的距离约为 30m。我们在原始全景图上执行数据平衡，以确保每个天线图像的正面全景不超过 2 幅（见图 3，分布包含在补充材料中）。该程序导致 105；214 幅全景图地理定位实验。此外，约 4% 的天线图像没有全景。我们把它们当作消遣以使数据集更真实和更具挑战性。卫星图像的缩放级别为 20，地面分辨率约为 0.114m。空中视图和地面视图的原始图像大小分别为 640×640 和 2048×1024。两种鸟瞰图的工业级 GPS 标签并且提供地面视图图像用于米级评估。然后根据方位信息变换全景图，使北方位于在中间。图 3显示了一对鸟瞰和街景图像。

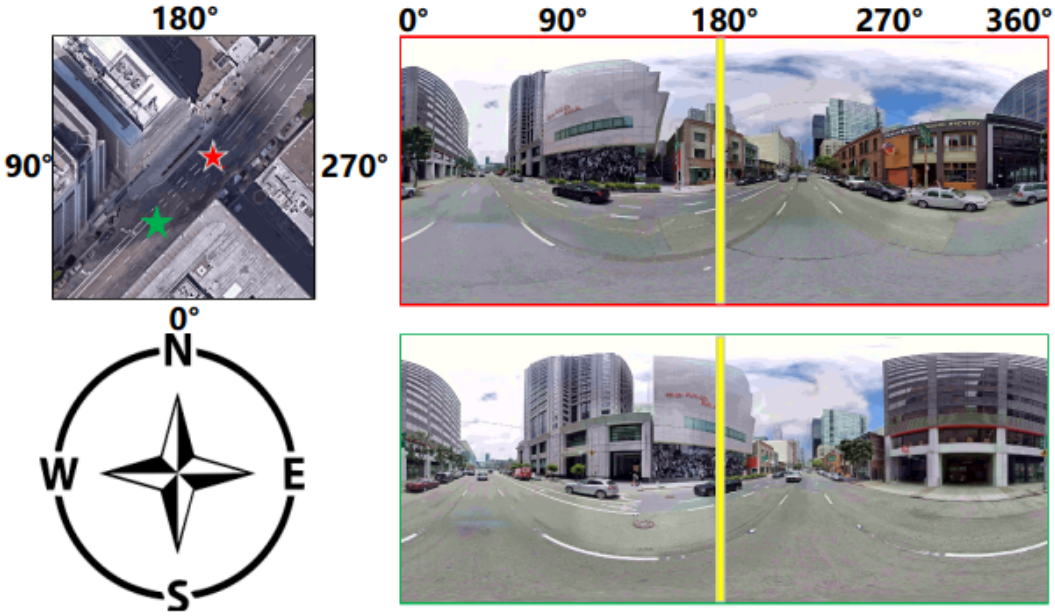

图 3: 正样本（恒星）和鸟瞰图与地面视图之间的方向对应的示例（黄色条形表示北）

头部对比：表 1显示了我们的数据集与以前的基准之间的比较。这个最广泛使用的数据集 CV - USA[11]由主要在郊区收集的图像组成。我们的数据集，位于另一方面，为城市环境收集数据。实际上，GPS 信号在城市地区比在郊区更容易产生噪音（例如，基于电话的 GPS 误差可能是在曼哈顿高达 50 米[1]）。因此，我们的数据集具有更多潜在的应用场景，例如基于视觉的移动辅助导航。此外，城市地区挤满了高楼。之间的相互语义遮挡和阴影显著减少了地面和鸟瞰图，使我们的数据集更具挑战性比 CVUSA。此外，以前的数据集仅采用用于评估的一对一检索，而不是因为无法预测并在那里预先捕

获对齐的参考图像。我们的数据集考虑了任意查询位置，甚至考虑了地面真相参考图像不具有与查询相同的 GPS 位置；因此，它更现实，但对检索具有挑战性。我们的数据集还提供了用于米级评估的原始 GPS 数据，这是定位应用的最终目标。我们相信我们的数据集是对现有的交叉视图图像数据集填补当前研究和实际应用。

	Vo [17]	CVACT [8]	CVUSA [19]	VIGOR (proposed)
Satellite images	∼ 450,000	128,334	44,416	90,618
Panoramas in total	∼ 450,000	128,334	44,416	238,696
Panoramas after balancing	-	-	-	105,214
Street-view GPS locations	Aligned	Aligned	Aligned	Arbitrary
Full panorama	No	Yes	Yes	Yes
Multiple cities	Yes	No	Yes	Yes
Orientation information	Yes	Yes	Yes	Yes
Evaluation in terms of meters	No	No	No	Yes
Seamless coverage on area of interest	No	No	No	Yes
Number of references covering each query	1	1	1	4

表 1: VIGOR 数据集与现有数据集之间的交叉视图地理定位比较[5,11-12]

评估方法: 根据不同的应用场景，我们为实验设计了两种评估设置，即相同区域和跨区域评估。相同区域：如果计划建立鸟瞰图参考 AOI 中任意街道查询的数据库，其目标是模型训练是处理任意的新查询。因此，最好的解决方案是收集 GPS 标记在同一区域进行培训而不是培训其他具有跨区域传输的区域。在这种情况下，天线四个城市的图像都作为参考数据用于培训和测试。然后所有的街道全景被随机分成两个不相交的集合（见表 2）。

		Same-Area		Cross-Area	
		Number	City	Number	City
Train	Aerial	90,618	All	44,055	New York
	Street	52,609	All	51,520	Seattle
Test	Aerial	90,618	All	46,563	San Francisco
	Street	52,605	All	53,694	Chicago

表 2: VIGOR 的评估分为两种设置

交叉区域：对于没有 GPS 标记查询的城市可用于训练，跨区域转移是必要的。在这种情况下，来自纽约和西雅图的所有图像用于训练，以及来自旧金山和芝加哥将等待评估。

3 从粗略到精细的横向视图本地化

在本章中，我们以粗略到精细的方式提出了用于地理定位的联合检索和校准框架。第 3.1 介绍了建立的强大基线采用最先进的技术样品。第 3.2 节提出了基于欠条的半正面分配损失以利用半阳性样本。使用检索到的最佳匹配参考图像，第 3.3 节旨在估计查询相对于检索到的中心的 GPS 位置作为仪表级校准的鸟瞰图像。

3.1 基准框架

为了在所提出的数据集上获得令人满意的结果，采用最先进的技术来建造强有力的底线。因此，我们采用 SAFA（空间感知特征聚合）的特征聚合模块[8-9]中的全球负面挖掘策略。功能聚合。SAFA[8]是极性变换、暹罗主干和特征聚合块的组合。然而，极坐标变换假设地面视图 GPS 位于相应的鸟瞰参考图像，不适用于我们的案例因此，我们只在我们的框架（见图 4）。功能的主要思想聚合块是根据嵌

入的位置重新加权嵌入。当定向时，空间感知块提供显著的性能增益查询图像的信息是可用的。挖掘策略：度量学习文献[15-17]已经揭示了在训练过程中挖掘硬样本的重要性，因为当大多数样本几乎没有贡献总损失时，模型会出现收敛性差的问题。对于横向地理定位，目前已经表明了开采全球硬样本而非开采的重要性在一个小批量内[9]。关键思想是构建一个先进先出的挖掘池，以缓存最硬样本的嵌入，并高效地刷新池以及反向传播。在小批量中，前一半图像是随机的所选的和关于每个的全局硬样本它们从采矿池中开采出来，形成另一半批次的。我们采用这种高效的全球采矿战略[9]以进一步提高其性能。

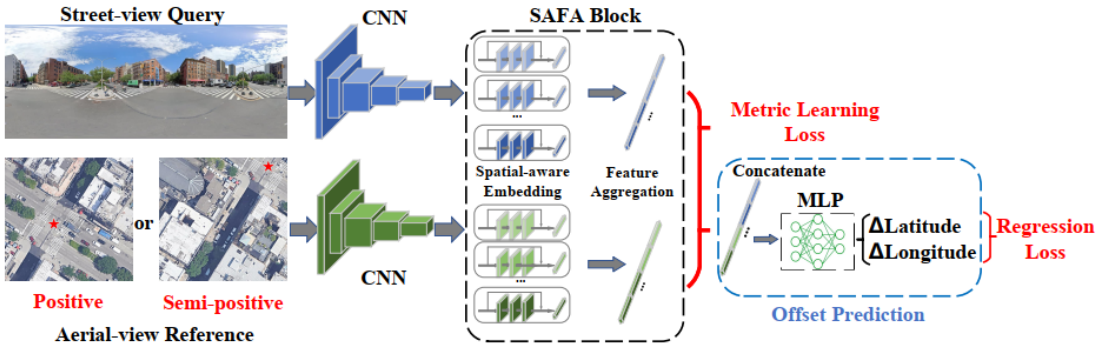

图 4: 网络模型图

3.2 基于半正样本的 iou

如果我们只考虑阳性样本这个问题可以通过标准度量学习来解决。对于我们采用了[7]中提出的广泛使用的三重态损失：

$$\mathcal{L}_{triplet} = log\left(1 + e^{\alpha(d_{pos} - d_{neg})}\right), \qquad (1)$$

其中 dpos 和 dneg 表示正对和负对。在一个具有 N 个地视图和鸟瞰图像对的小批量中，我们使用策略[15]构建 2N（N−1）个三联体，从而充分使用所有输入图像。根据[7]，我们采用 l2 对输出嵌入特征进行归一化。

3.3 偏移量预测

对于前 1 个检索到的参考航空图像，我们使用一个辅助任务来在统一的框架中进一步细化航空视图图像内部的定位（见图 4). 对于图像检索，我们数据集中检索到的参考图像之间的最小间隔是宽度的一半航空图像（L=2）。为了实现更细粒度的定位，我们应用 MLP（多层感知器）来预测查询位置相对于检索到的参考图像。如图 4所示，辅助 MLP 由两个完全连接的层组成连接的嵌入特征作为输入。这里我们使用回归生成预测，同时我们还提供与实验中的分类进行比较。这个补偿回归损失公式为：

$$\mathcal{L}_{offset} = (lat - lat^*)^2 + (lon - lon^*)^2,$$

其中 lat 和 lon 表示查询的预测偏移量相对于参考 GPS 的 GPS 位置（lat*、lon* 表示地面实况偏移。它们都被转换成米，并在训练期间用 L 标准化。最终混合损失函数由下式给出：

$$\mathcal{L}_{hybrid} = \mathcal{L}_{triplet} + \mathcal{L}_{IOU} + \mathcal{L}_{offset}.$$

4 实验

4.1 实施细节

所有实验都已展开基于 Tensorflow[18]。地面全景和鸟瞰图像大小调整为 640×320 和 320×320 分别在被馈送到网络之前。VGG-16[19] 被用作主干特征提取器，并且通过以下[8]使用 8 个 SAFA 块。采矿策略参数的设置与[9]中的相同。继[7]之后，我们将 Ltriplet 损失中的 α 设置为 10。Adam 优化器[20]学习率为 10−5。我们的方法经过培训对于相同区域设置，45 个时期交叉区域设置。用于比较的基准（第 4.1 节）仅使用 Ltriplet 进行训练。由于作者提出的跨视角图像检索方法已经取得了较好的检测效果，因此这里提出了新的思想，即探究图像的尺寸和特征向量维度的关系。为了更好的研究，本文对 SAFA 模块进行了修改，修改的思想如图 5所示。

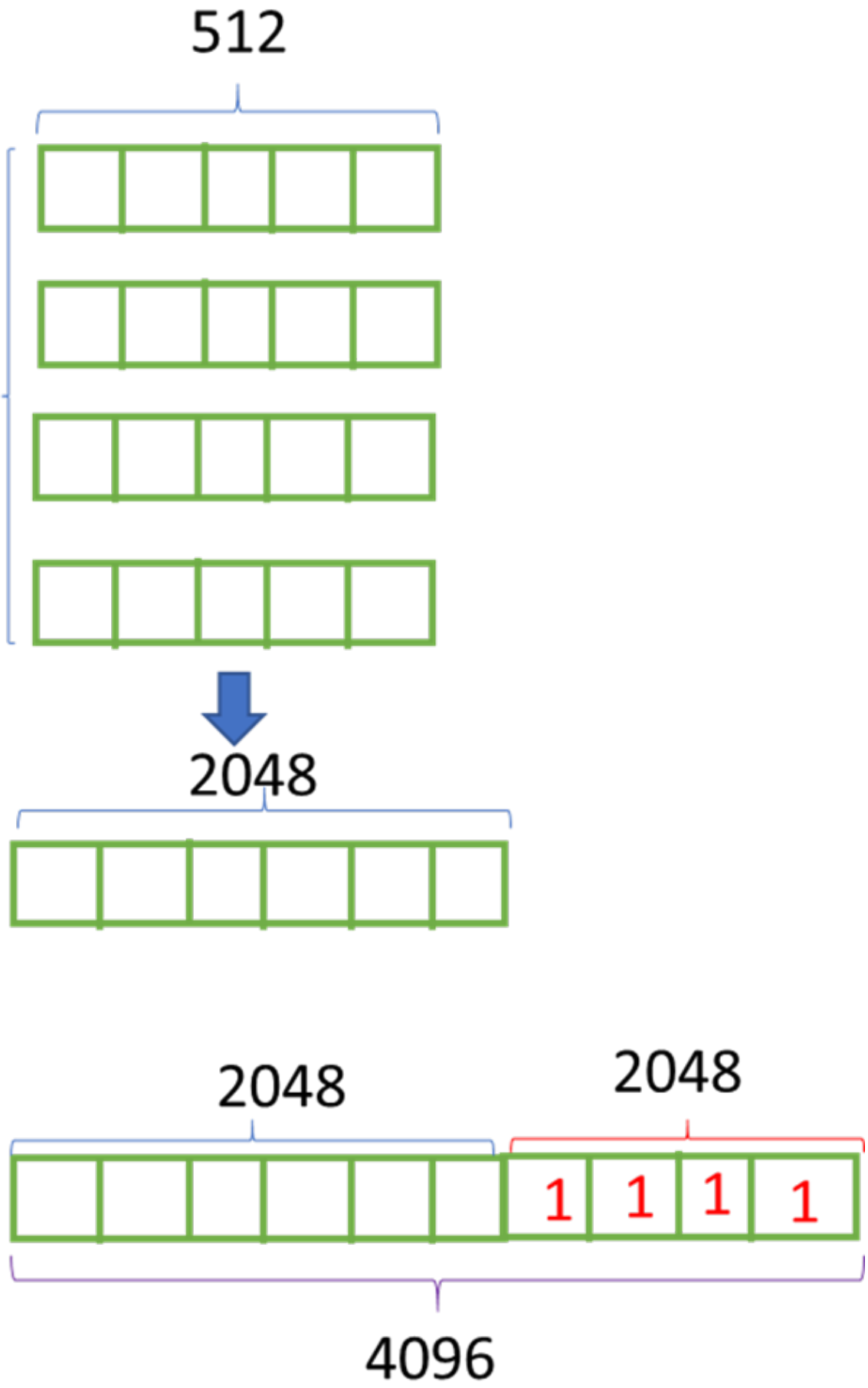

图 5: SAFA 模块修改思想

评估指标：我们首先评估了检索性能，在前面的步骤中，我们使用了前 k 个召回准确率工作[8]。对

于每个测试查询，其最近的 k 个引用检索嵌入空间中的邻居作为预测。如果基本事实属实，则认为一次检索是正确的图像被包括在前 k 个检索到的图像中。如果检索到的 top 1 参考图像覆盖了查询图像（包括地面真相），则它被视为命还为检索评估提供了命中率。此外我们计算了排名前 1 的预测位置和地面实况查询 GPS 之间的真实世界距离，作为米级评估.

4.2 实验结果

图 6为作者在论文给出的模型学习后的最终结果，为了减少对比实验的数量，这里仅对 SAFA 模块的网络在相同面积的情况下进行实验，从表中可以看出，top1 的精度为 33.9%，top5 的精度为 58.4%,top-1% 的精度为 98.2%,hit rate 的精度为 36.9%。

	Same-Area				Cross-Area			
	Top-1	Top-5	Top-1%	Hit Rate	Top-1	Top-5	Top-1%	Hit Rate
Siamese-VGG ($\mathcal{L}_{triplet}$)	18.1	42.5	97.5	21.2	2.7	8.2	61.7	3.1
SAFA ($\mathcal{L}_{triplet}$)	33.9	58.4	98.2	36.9	8.2	19.6	77.6	8.9
SAFA+Mining (baseline, $\mathcal{L}_{triplet}$)	38.0	62.9	97.6	41.8	9.2	21.1	77.8	9.9
Ours ($\mathcal{L}_{hybrid}$)	**41.1**	**65.8**	**98.4**	**44.7**	**11.0**	**23.6**	**80.2**	**11.6**

Table 3. Retrieval accuracy (percentage) of different methods.

图 6: 论文中给出的模型结果）

图 7是未对论文中的模型做任何修改时，模型最终得到的结果。从结果中可以看出，top1 的精度为 29.1%，top5 的精度为 53.3%,top-1% 的精度为 97.6%,hit rate 的精度为 31.5%。对比图 6可知，各项指标的整体精度均有所下降，但是 top-1% 的精度几乎与原有精度下降较小。图 8为最终得到的特征向量在不同阈值情况下的精度，从下图可以看出，随着阈值的增大，精度逐渐提高。

epoch	top-1%	top-1	top-5	hit rate
1	82.0%	4.1%	11.6%	4.3%
2	88.7%	6.9%	17.5%	7.2%
3	88.6%	6.9%	17.7%	7.2%
4	93.5%	11.3%	26.0%	11.9%
5	94.7%	13.3%	29.8%	14.1%
6	95.7%	15.3%	33.3%	16.2%
7	95.4%	15.3%	32.6%	16.2%
8	96.1%	17.3%	36.2%	18.4%
9	95.9%	17.3%	35.9%	18.3%
10	96.9%	20.0%	40.4%	21.4%
11	96.9%	20.0%	40.6%	21.2%
12	97.2%	22.1%	43.6%	23.6%
13	97.2%	22.6%	44.0%	24.1%
14	97.3%	22.7%	44.3%	24.3%
15	97.3%	22.4%	44.1%	24.0%
16	97.1%	24.5	46.7%	26.3%
17	97.3%	23.7%	46.0%	25.6%
18	97.6%	25.6%	48.6%	27.6%
19	97.5%	26.0%	49.0%	28.0%
20	97.2%	24.2%	46.9%	26.2%
21	97.8%	27.0%	50.1%	29.1%
22	97.7%	27.1%	50.7%	29.3%
23	97.7%	28.0%	51.6%	30.1%
24	97.7%	26.6%	50.0%	28.8%
25	97.7%	27.5%	51.2%	29.6%
26	97.5%	26.5%	49.7%	28.6%
27	97.7%	28.4%	52.3%	30.7%
28	97.6%	29.6%	53.6%	32.1%
29	97.6%	29.1%	53.2%	31.6%
30	97.6%	29.1%	53.3%	31.5%

图 7: 论文中原始模型的结果

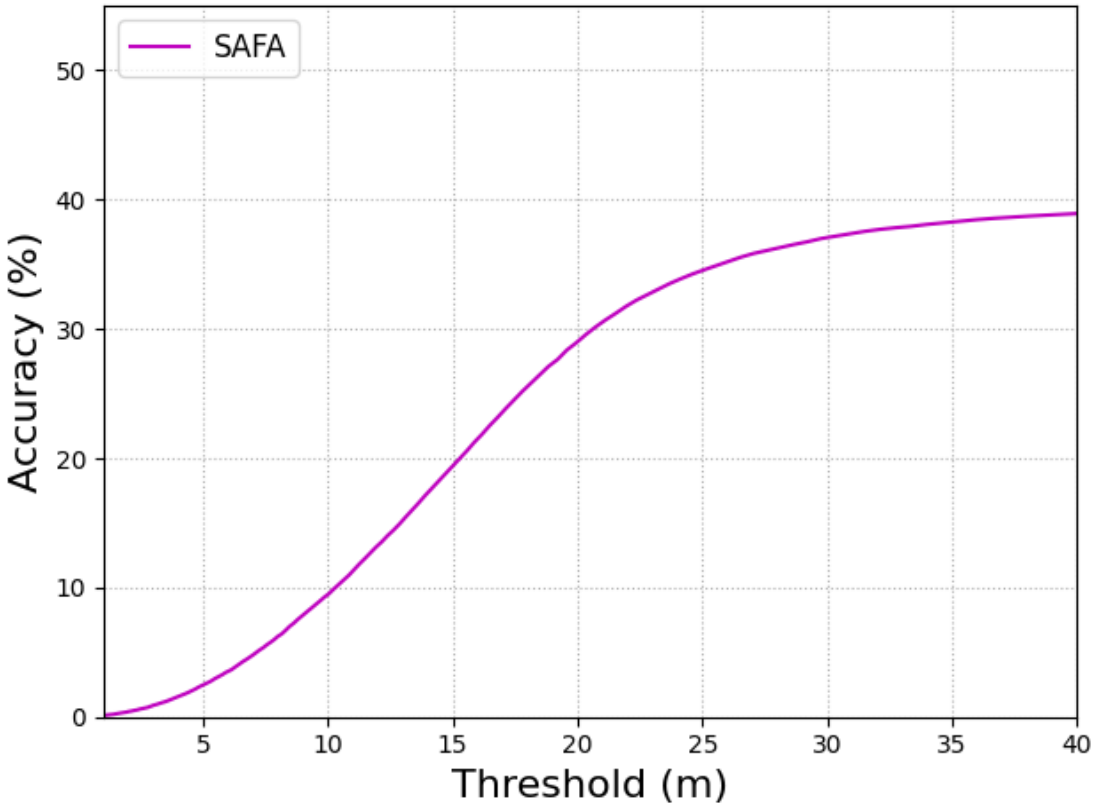

图 8: 阈值与精度的关系图

图 9为修改 SAFA 模块后的结果，从图中可以看出，top1 的精度为 29.0%，top5 的精度为 52.8%,top-1% 的精度为 97.7%,hit rate 的精度为 31.6%。与论文中提到的模型检测结果相比，整体精度几乎没有变化。图 10为改变模型后的阈值与精度的关系图，对比图 8可知，二者几乎没有太大变化。综上，即按着本文提出的思路改变图像的特征向量维度，对检测结果不会带来太大影响。

epoch	top-1%	top-1	top-5	hit rate
1	79.5%	3.5%	9.9%	3.7%
2	89.3%	6.8%	17.4%	7.1%
3	91.5%	8.3%	20.7%	8.8%
4	93.0%	10.4%	24.6%	10.9%
5	94.5%	12.2%	28.2%	12.9%
6	95.5%	14.6%	32.2%	15.5%
7	95.7%	15.3%	33.8%	16.2%
8	96.2%	16.8%	35.8%	17.9%
9	96.3%	16.8%	36.2%	18.0%
10	97.0%	18.4%	38.7%	19.7%
11	96.8%	19.0%	39.3%	20.4%
12	96.9%	19.4%	40.6%	20.9%
13	97.1	21.7%	43.0%	23.1%
14	97.2%	22.0%	43.9%	23.6%
15	97.2%	22.7%	44.6%	24.5%
16	97.2%	22.6%	44.3%	24.2%
17	97.5%	23.6%	46.1%	25.5%
18	97.6%	24.1%	47.0%	26.0%
19	97.4%	23.9%	46.4%	25.7%
20	97.4%	23.1%	45.7%	25.1%
21	97.5%	25.1%	48.2%	27.1%
22	97.4%	25.8%	49.0%	27.8%
23	97.6%	26.5%	50.0%	28.6%
24	97.9%	27.2%	50.9%	29.6%
25	97.5%	26.4%	49.8%	28.7%
26	97.8%	27.6%	51.3%	29.8%
27	97.7%	27.6%	51.8%	30.0%
28	97.3%	26.3%	49.7%	28.5%
29	97.7%	28.4%	52.5%	31.0%
30	97.7%	29.0%	52.8%	31.6%

图 9: 改变 SAFA 模块后的结果

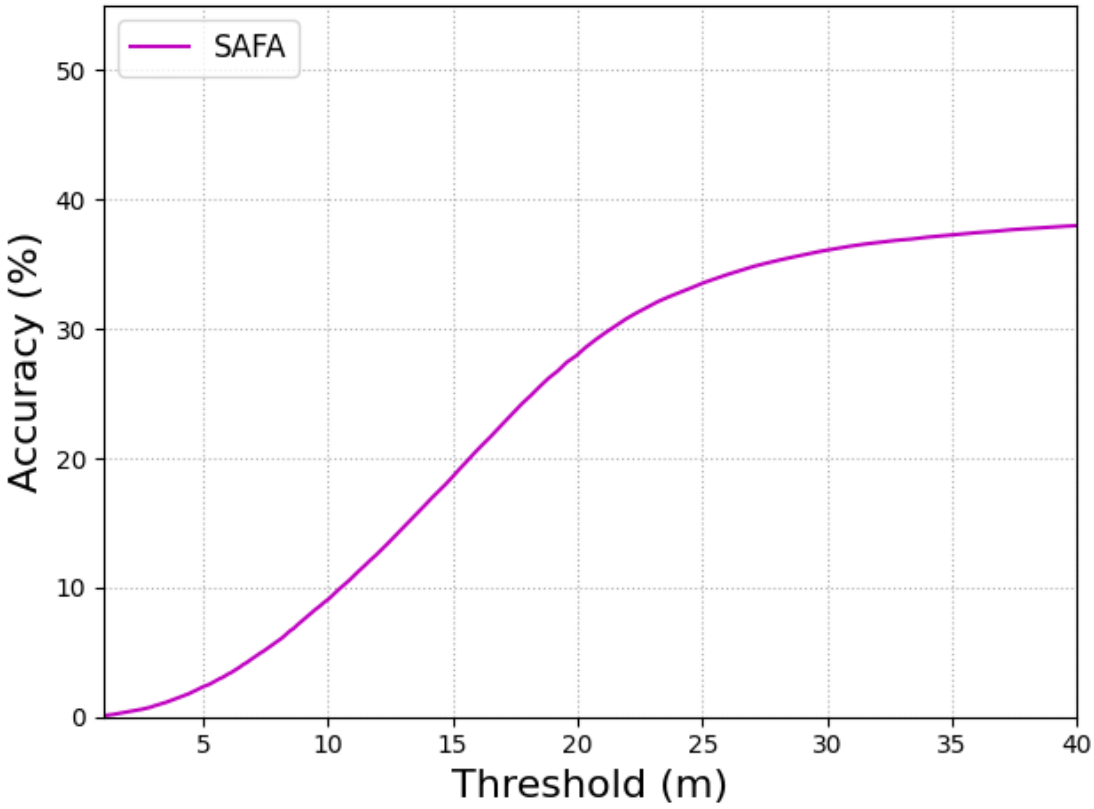

图 10: 改变 SAFA 后的阈值与精度的关系图

5 总结与展望

通过此次研究，可以得到以下结论和研究内容：（1）本文复现了作者提出的实验，但是精度尚未达到论文中提到的精度，具体原因尚未弄明白，还需进一步研究。（2）初步论证了本文提出的修改方法基本不会影响模型的精度。（3）由于维度几乎不影响检测的精度，因此，推断现有的 4096 维度向量是高度冗余的。同时，也为进行跨视角目标检测提供了数据支持。

参考文献

[1] BROSH E, FRIEDMANN M, KADAR I, et al. Accurate visual localization for automotive applications [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2019: 0–0.

[2] ZAMIR A R, SHAH M. Accurate image localization based on google maps street view[C]//European Conference on Computer Vision. 2010: 255-268.

[3] MIROWSKI P, GRIMES M, MALINOWSKI M, et al. Learning to navigate in cities without a map[J]. Advances in neural information processing systems, 2018, 31.

[4] LI A, HU H, MIROWSKI P, et al. Cross-view policy learning for street navigation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 8100-8109.

[5] VO N N, HAYS J. Localizing and orienting street views using overhead imagery[C]//European conference on computer vision. 2016: 494-509.

[6] TIAN Y, CHEN C, SHAH M. Cross-view image matching for geo-localization in urban environments

[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 3608-3616.

[7] HU S, FENG M, NGUYEN R M, et al. Cvm-net: Cross-view matching network for image-based ground-to-aerial geo-localization[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7258-7267.

[8] SHI Y, LIU L, YU X, et al. Spatial-aware feature aggregation for image based cross-view geo-localization[J]. Advances in Neural Information Processing Systems, 2019, 32.

[9] ZHU S, YANG T, CHEN C. Revisiting street-to-aerial view image geo-localization and orientation estimation[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021: 756-765.

[10] SUN B, CHEN C, ZHU Y, et al. GEOCAPSNET: Ground to aerial view image geo-localization using capsule network[C]//2019 IEEE International Conference on Multimedia and Expo (ICME). 2019: 742-747.

[11] ZHAI M, BESSINGER Z, WORKMAN S, et al. Predicting ground-level scene layout from aerial imagery[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 867-875.

[12] LIU L, LI H. Lending orientation to neural networks for cross-view geo-localization[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 5624-5633.

[13] https://developers.google.com/maps/documentation/mapsstatic/intro.[J].,

[14] https://developers. streetview/intro.goo-gle.com/maps/documentation/streetview/intro[J].,

[15] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A unified embedding for face recognition and clustering[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 815-823.

[16] SUH Y, HAN B, KIM W, et al. Stochastic class-based hard example mining for deep metric learning [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 7251-7259.

[17] MUSGRAVE K, BELONGIE S, LIM S N. A metric learning reality check[C]//European Conference on Computer Vision. 2020: 681-699.

[18] ABADI M, BARHAM P, CHEN J, et al. {TensorFlow}: a system for {Large-Scale} machine learning [C]//12th USENIX symposium on operating systems design and implementation (OSDI 16). 2016: 265-283.

[19] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv:1409.1556, 2014.

[20] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.