arXiv:2303.02058v2 [cs.CV] 2 Aug 2023

利用高斯隐式占据函数进行3D感知目标定位

Vincent Gaudilli`ere Leo Pauly Arunkumar Rathinam Albert Garcia Sanchez Mohamed Adel Musallam DjamilaAouada

摘要——**在图像中自动定位目标物体对许多计算机视觉应用至关重要。为了表示二维物体，近期研究发现，椭圆标签是轴对齐边界框的一种有前景的替代方案。本文进一步针对二维目标定位考虑3D感知椭圆标签，即物体的三维椭球近似在图像上的投影椭圆。事实上，与传统的不考虑三维信息的边界框标签相比，投影椭圆携带了更多关于物体几何形状和姿态（三维感知）的信息。此外，这种通用的三维椭球模型能够用于近似已知或仅粗略已知的目标。随后，我们重新审视椭圆回归，并用隐式高斯分布的参数取代不连续的椭圆几何参数，以此对图像中的物体占据情况进行编码。模型经过训练，可利用统计损失函数回归该二元高斯分布在图像像素上的取值。我们提出一种新颖的非可训练可微层 E-DSNT，用于提取分布参数。此外，我们还介绍了如何仅根据目标的粗略尺寸和相对姿态标签，直接生成一致的3D感知高斯占据参数。我们为三个现有的航天器姿态估计数据集扩充了3D感知高斯占据标签，以验证我们的假设。标签和源代码可在此处公开获取：https://cvi2.uni.lu/3d-aware-obj-loc/。**

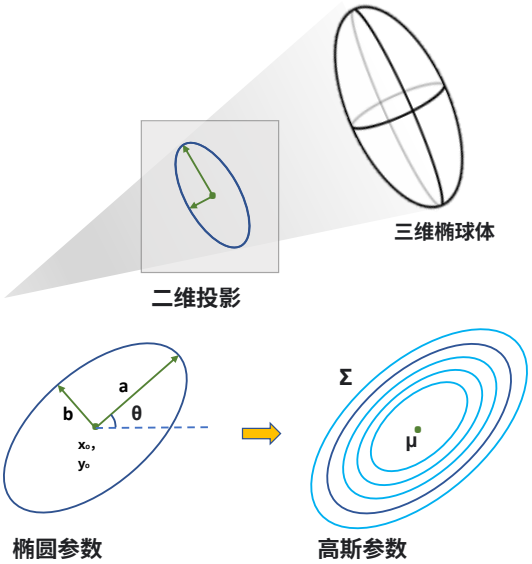


图1。**上图：**我们考虑三维椭球体的二维投影，以近似表示该物体。**左下图：**标准椭圆表示，包含中心坐标（x_0，y_0）、长轴（a）和短轴（b），以及方向θ。**右下图：**为处理 θ的不连续性，我们提出用一个底层隐式高斯分布来表示椭圆，该分布由连续的均值 $\mu = x_0\ y_0$ 和协方差 $\Sigma = (,)$ cos( ) sin( ) sin( ) cos( ) cos( ) sin( ) sin( ) cos( ) 参数化。

$$\begin{pmatrix} \theta & -\theta \\ \theta & \theta \end{pmatrix} \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix} \begin{pmatrix} \theta & \theta \\ -\theta & \theta \end{pmatrix}$$

I. 引言

图像中的目标定位因其有望对广泛应用产生影响而受到计算机视觉领域的关注。尽管轴对齐边界框一直是目标检测事实上的标准表示形式， [1]，椭圆最近被认为是另一种通用表示，能够承载更多有关物体投影的信息，例如其方向和更贴合的包络，从而例如实现更精确的三维重建 [2]。

二维椭圆回归 椭圆回归的开创性工作一直聚焦于椭圆几何参数，即中心坐标、短轴和长轴以及方向角度（见图1）。为了回归这些参数，椭圆提议网络 [3] 使用 L1 损失，因此需要在位置和轴误差与方向误差之间进行相对加权。为避免这一问题，Lin等人 [4] 对这些参数进行变换，使其处于相同的数量级。这些方法的主要局限在于待回归角度值的不连续性。

这项工作部分由卢森堡国家研究基金（FNR）资助，项目编号为 BRIDGES2020/IS/14755859/MEET- A/Aouada。

作者来自卢森堡大学 SnT——跨学科安全、可靠性与信任中心。联系邮箱：vincent.gaudilliere@uni.lu

高斯提议网络（GPN） [5]，受区域提议网络（RPN） [6],启发，将椭圆视为图像平面上的二维高斯分布，并将提议分布与真实分布之间的 Kullback-Leibler（KL）散度作为单一损失函数进行最小化。由于高斯分布的 KL 散度具有解析形式且可微，GPN 易于实现，并可通过反向传播算法进行训练。然而，高斯参数是从回归得到的几何参数（位置、轴和方向）推导而来，因此也面临角度回归中的同一不连续性问题。Pan 等人 [7]，将 GPN 集成到目标检测流水线中，并使用 Wasserstein 度量代替 KL 散度，为模型提供适当的距离损失。Ellipse R-CNN [2]也回归椭圆几何参数，并增强了遮挡鲁棒性；而 ElDet [8] 则增加了训练目标，例如最大化预测椭圆与真实椭圆之间的交并比（IoU）得分。然而，这些方法旨在检测图像中自然存在的

椭圆形状，这与我们的应用场景并不相符；在我们的场景中，椭圆源自用于建模物体的三维虚拟椭球体。

三维感知椭圆回归据我们所知，Zins 等人 [9], [10], [11] 首次对三维感知椭圆进行了回归。在 [9], 中，他们对椭圆中心和尺寸采用 L2 损失，而角度预测则被建模为带有后验角度校正损失的分类问题。他们还提出了两类由椭圆参数表征的隐式函数，以增强损失函数的鲁棒性：局部有符号距离函数，可实现真实值与预测函数值之间的逐像素比较 [10]；以及基于椭圆方程的代数距离函数，结合自适应采样以实现旋转不变性 [11]。然而，隐函数值仍然是基于回归得到的椭圆几何参数计算的。

本文沿用这一隐函数思想，但提出以占用热图的形式，在图像像素上连续回归其数值。训练前，利用相对物体—相机位姿，直接根据椭球投影计算高斯分布参数和热力图。在前向传播过程中，我们借助一种新颖的非训练式可微层——扩展可微空间到数值变换（E-DSNT），从回归的热力图中提取这些参数。最后，提出结合多种统计损失来优化模型。尽管我们的方法可用于任何需要二维目标定位的应用场景，本文重点关注空间态势感知（SSA）应用，以对其进行验证。实际上，自动定位非合作航天器目标对于在轨交会等任务至关重要。我们在三个航天器位姿估计基准数据集上评估了本方法。简而言之，我们的贡献有以下三点：

- 一种新颖且完全可微的物体定位流程，可直接从图像回归3D感知椭圆标签。该方法在标准航天器定位基准测试中取得了最先进性能；
- 一种仅根据六自由度相对位姿和粗略物体尺寸生成3D感知高斯占据标签的方法；
- 开放获取发布适用于三个现有航天器姿态估计数据集的3D感知高斯占据标签（热力图、均值与协方差标签）。

本文其余部分的组织结构如下。第二节介绍用于图像中物体定位的3D感知高斯占据标签的生成方法。第三节介绍我们设计的、用于回归此类标签的物体定位模型。随后，第四节给出实验对比，展示该方法的最先进性能。最后，第五节总结全文。

II. 3D感知高斯占据标签

本节介绍如何根据物体的三维椭球模型轻松生成高斯占据标签



图2. 边界框对象标签（左）和椭圆标签（右）的示意图。对于许多人形物体，椭圆是更准确的表示。图像来自 URSO 数据集 [12]。

以及相对相机—物体位姿。

A. 二维物体占据

图像中的物体占据可定义为与该物体对应的像素集合。在大多数情况下，目标检测标签 L_{obj} 采用边界框的形式，旨在包围物体占据区域（见图 2）。对于中心为 $(x_c,\ y_c)$、尺寸为 $(W,\ H)$ 的边界框，目标检测标签可表示为：

$$L_{obj}^{(bbox)} = [x_c, y_c, W, H]\,. \tag{1}$$

现在假设物体占据由隐式占据函数 $O_{\mathbf{p}} : \mathbb{R}^2 \to \mathbb{R}$ 表示，其由参数$\mathbf{p}$表征，则物体标签可写为

$$L_{obj} = \mathbf{p}. \tag{2}$$

B. 二维高斯占据标签

物体占据的统计信息由二维图像中被占据像素的均值 μ_x，μ_y、方差 Σ_{xx}，Σ_{yy}和协方差 Σ_{xy}给出。因此，一种自然的隐式占据函数是二元高斯分布，其中心为 $\mu = (\mu_x,\ \mu_y)^\top$，协方差矩阵为 $\Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy} & \Sigma_{yy} \end{pmatrix}$。最后，我们将高斯占据标签定义为

$$L_{obj}^{(Gauss)} = [\mu_x, \mu_y, \Sigma_{xx}, \Sigma_{yy}, \Sigma_{xy}]. \tag{3}$$

C. 3D感知高斯占据标签

二元高斯分布的参数与特征椭圆的参数相同，反之亦然。在我们的方法中，我们关注的是回归三维椭球体的椭圆投影参数，因此考虑由其产生的高斯分布。

更准确地说，给定一个物体，其沿三个正交主方向的尺寸分别为 $2a'$、$2b'$、$2c'$，该物体的椭球近似由矩阵表示

$$Q^* = diag\begin{pmatrix} a'^2 & b'^2 & c'^2 & -1 \end{pmatrix}. \tag{4}$$

接着，令 T 表示物体相对于相机的中心位置，R 表示物体朝向，则椭球在图像中的投影（椭圆）可由矩阵表示：

$$C^* = PQ^*P^\top, \tag{5}$$

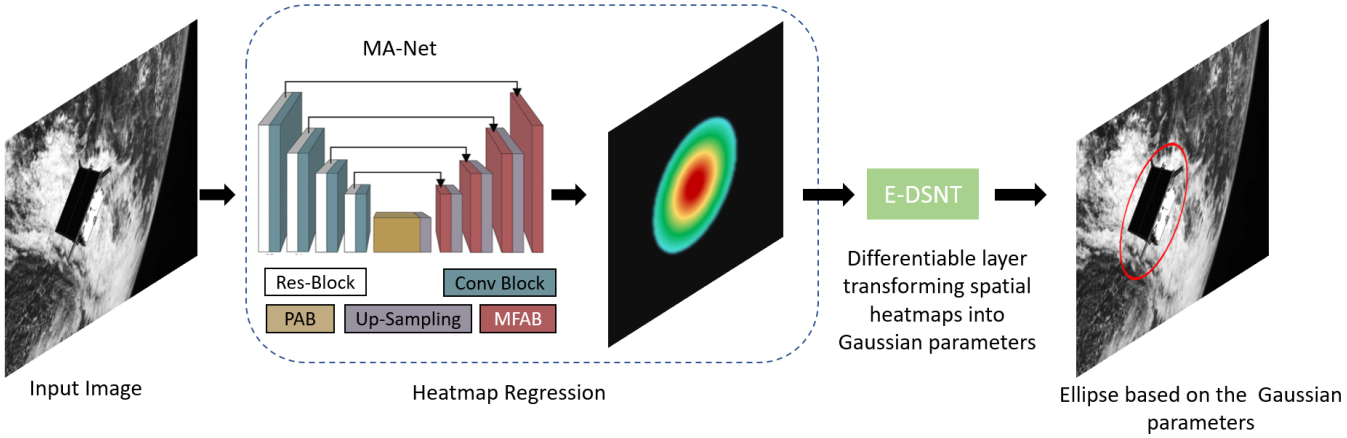


图 3. 我们的物体定位流程示意图，旨在以完全可微的方式回归三维感知高斯隐式占据函数的参数（在中央图中以热力图表示）。该模型结合了 MA-Net 骨干网络，用于回归图像像素上的隐函数值；并采用一种新颖的非训练型可微层（E-DSNT）来提取高斯分布参数。随后得到相应的椭圆（叠加在右图中，以红色显示）。图像来自 SPEED+ 数据集 [13]。

其中 $P = \mathrm{K}\,[\mathrm{R}|\mathrm{T}]$，K 为相机内参矩阵 [14]。因此，投影椭圆 C^* 的形式为：

$$C^* = s \begin{pmatrix} \cdot & \cdot & -x_0 \\ \cdot & \cdot & -y_0 \\ -x_0 & -y_0 & -1 \end{pmatrix}, \tag{6}$$

其中，s 是缩放因子，$\cdot$ 表示任意值，而 $(x_0,\ y_0)$ 是椭圆中心，对应于底层高斯占据函数的峰值位置：

$$\mu = (\mu_x, \mu_y)^\top = (x_0, y_0)^\top. \tag{7}$$

随后可得以该中心为中心的椭圆：

$$C^*_{\text{centered}} = TC^*T^\top, \tag{8}$$

where $T = \begin{pmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{pmatrix}$.

2×2 的左上部分 C^*_{centered} 对应于高斯占据函数的协方差 Σ，其特征分解可得椭圆方向 θ 和半轴（$a,\ b$）：

$$\Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy} & \Sigma_{yy} \end{pmatrix} \tag{9}$$

$$= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{10}$$

III. 目标定位

如图 3 所示，我们的物体定位流程可以通过隐式函数估计，以完全可微的方式回归高斯占据标签 $L_{obj}^{(Gauss)}$。我们引入了一个可微的热力图参数提取层 E-DSNT，并将其与热力图回归网络结合，以估计隐函数值。这些模块分别在第三节 B 和第三节 A 中介绍。用于优化模型的统计损失在第三节 C 中介绍。

A. 热力图回归网络

任何最先进的分割网络都可用于回归图像像素上的高斯占据函数值 $O_{\mathbf{p}}(x,\ y)$。在实验中，我们使用 MA-Net [15] 网络及其 ResNet34 [16]骨干网络，因为该模型能够通过注意力机制捕获上下文依赖关系，并采用两个模块（见图 3）：位置注意力模块（PAB）用于从全局视角捕获像素之间的空间依赖关系；多尺度融合注意力模块（MFAB）则通过多尺度语义特征融合，捕获任意特征图之间的通道依赖关系。最后，添加一个展平-Softmax 层，以确保得到归一化的概率密度函数，以下称之为热力图。

B. 高斯占据标签的可微提取

二元高斯分布由其均值 μ（即峰值的坐标）以及协方差矩阵 Σ（编码分布的空间范围）表征。文献中此前已介绍过一个可微均值提取层（见第三节-B.1），我们将其扩展为提取额外的协方差矩阵值，详见第三节-B.2。

1) 均值提取：可微空间到数值变换（DSNT）[17]是一种不可训练且可微的层，用于从给定的、尺寸为 $H \times W$ 的归一化˜高斯热力图 Z 中提取均值。按照此方法，我们计算两个编码坐标的矩阵 X和 Y，其元素为

$$X_{i,j} = \frac{2j - (W+1)}{W};\ Y_{i,j} = \frac{2i - (H+1)}{H},$$
$$\forall i = 1,\ \ldots,\ H;\ \forall j = 1,\ \ldots,\ W.$$

注意到热力图编码了像素 $[X_{i,j}\ , Y_{i,j}]$ 成为峰值 p 所在位置的概率 P，我们有：

$$P(p = [X_{i,j}\ , Y_{i,j}]) = \hat{Z}_{i,j}.$$

p 的预测通过其期望得出

$$\mu = \mathbb{E}[p] = \begin{bmatrix} < \hat{Z}, X >_F & < \hat{Z}, Y >_F \end{bmatrix} \tag{11}$$

其中，$< \cdot, \cdot >_F$ 是 Frobenius 内积。

2）协方差矩阵提取：令 p_x，p_y 表示 p的坐标，我们扩展 DSNT，以提取热力图在 x 轴 Σ_{xx} 和 y 轴 Σ_{yy} 上的方差，以及协方差值 Σ_{xy}。我们将此类参数提取过程称为扩展-DSNT（E-DSNT）。具体而言，我们使用上述量的定义，并推导出以下公式：

$$\begin{aligned} \Sigma_{xx} &= \mathbb{E}[(p_x - \mathbb{E}[p_x])^2] \\ &=< \hat{Z}, (X - \mu_x) \odot (X - \mu_x) >_F \end{aligned} \tag{12}$$

$$\begin{aligned} \Sigma_{yy} &= \mathbb{E}[(p_y - \mathbb{E}[p_y])^2] \\ &=< \hat{Z}, (Y - \mu_y) \odot (Y - \mu_y) >_F \end{aligned} \tag{13}$$

$$\begin{aligned} \Sigma_{xy} &= \mathbb{E}[(p_x - \mathbb{E}[p_x])(p_y - \mathbb{E}[p_y])] \\ &=< \hat{Z}, (X - \mu_x) \odot (Y - \mu_y) >_F \end{aligned} \tag{14}$$

因此，高斯分布参数包括中心 μ（由公式（11）得到），以及由公式（12）、（13）和（14）左侧各项构成的协方差矩阵 Σ。

3）椭圆参数计算：椭圆的中心即为（x_0，y_0）$= \mu$。其轴 a，b和方向 θ 通过对协方差矩阵 Σ 进行特征分解得到（见公式 10）。

C. 模型损失

我们的损失 $\mathcal{L}$ 可表示为两种损失的组合。我们使用 Wasserstein 距离 $\mathcal{L}_{\mathrm{W}}$ 直接优化高斯参数，同时使用作用于热图值的 Jensen-Shannon 散度 $\mathcal{L}_{\mathrm{JS}}$来正则化隐式占据函数 $O_{\mathbf{p}}$。因此，我们的损失为

$$\mathcal{L} = \mathcal{L}_{\mathrm{W}} + \lambda . \mathcal{L}_{\mathrm{JS}}, \tag{15}$$

并使用标量因子 λ 来平衡这两项损失。

具体而言，考虑由均值和协方差（μ_{pd}，Σ_{pd}）和（μ_{gt}，Σ_{gt}）表征的预测高斯分布 $\mathcal{G}_{\mathrm{pd}}$ 与真实高斯分布 $\mathcal{G}_{\mathrm{gt}}$，Wasserstein 距离项定义如下：

$$\begin{aligned} \mathcal{L}_{\mathrm{W}} &= d_W(\mathcal{G}_{\mathrm{pd}}, \mathcal{G}_{\mathrm{gt}}) \\ &= ||\mu_{\mathrm{pd}} - \mu_{\mathrm{gt}}||_2^2 + \mathrm{tr}\left(\Sigma_{\mathrm{pd}} + \Sigma_{\mathrm{gt}} - 2\left(\Sigma_{\mathrm{gt}}^{\frac{1}{2}}\Sigma_{\mathrm{pd}}\Sigma_{\mathrm{gt}}^{\frac{1}{2}}\right)^{\frac{1}{2}}\right). \end{aligned} \tag{16}$$

这种闭式表达式无需依赖手工设定的相对权重来平衡均值、方差项和协方差项的贡献。

Jensen-Shannon 散度项定义如下：

$$\mathcal{L}_{\mathrm{JS}} = \frac{1}{2} D_{\mathrm{KL}}(\mathcal{G}_{\mathrm{pd}} \| \mathcal{G}_{\mathrm{m}}) + \frac{1}{2} D_{\mathrm{KL}}(\mathcal{G}_{\mathrm{gt}} \| \mathcal{G}_{\mathrm{m}}), \tag{17}$$

其中，$\mathcal{G}_{\mathrm{m}} = \frac{1}{2}(\mathcal{G}_{\mathrm{pd}} + \mathcal{G}_{\mathrm{gt}})$ 和 $D_{\mathrm{KL}}(\mathcal{D}_1 \| \mathcal{D}_2)$ 是任意分布 $\mathcal{D}_1$与 $\mathcal{D}_2$ 之间的 Kullback-Leibler 散度，定义为

$$D_{\mathrm{KL}}(\mathcal{D}_1 \| \mathcal{D}_2) = \sum_{(i,j) \in I} \mathcal{D}_1(i,j) \log \left(\frac{\mathcal{D}_1(i,j)}{\mathcal{D}_2(i,j)} \right). \tag{18}$$

该项直接根据热力图 I中各像素（i，j）处的隐函数值计算。Jensen-Shannon 散度具有对称这一优点，不同于 Kullback-Leibler；且已有研究证明，其表现优于后者 [17]。

IV. 实验与讨论

扩展数据集并添加具备 3D 感知能力的高斯占据标签我们为三个公开的航天器姿态估计数据集添加了高斯占据标签。SPEED [20]和SPEED+ [13] 是航天器姿态估计方法的标准基准，而 AKM [21]是近期发布的数据集，包含无纹理且对称的空间物体 [22]。前两个数据集均不提供 CAD数据，但我们的方法只需要所考虑航天器的大致尺寸（TANGO 尺寸：80×75×32cm [23]）。

目标定位指标 我们的指标包括交并比（记为 IoU）、重叠度（Overlap）、Dice-Sorensen系数（Dice）、相对体积差（RVD）和改进豪斯多夫距离（MHD） [24]。令 P 和 G 分别表示预测椭圆和真实椭圆内部的像素集合，这些指标定义如下：

$$\mathrm{IoU} = \frac{\mathrm{area}(P \cap G)}{\mathrm{area}(P \cup G)}, \tag{19}$$

$$\mathrm{Overlap} = \frac{\mathrm{area}(P \cap G)}{\min(\mathrm{area}(P), \mathrm{area}(G))}, \tag{20}$$

$$\mathrm{Dice} = 2\frac{\mathrm{area}(P \cap G)}{\mathrm{area}(P) + \mathrm{area}(G)}, \tag{21}$$

$$\mathrm{RVD} = \frac{|\mathrm{area}(P) - \mathrm{area}(G)|}{\mathrm{area}(G)}. \tag{22}$$

这四项指标从不同角度刻画了由预测椭圆和真实椭圆围成的区域 P和 G 之间的差异，而 MHD [24]则直接衡量（离散）椭圆 ∂P 和 ∂G之间的差异：

$$\mathrm{MHD} = \max\left(\mathrm{mhd}(\partial P, \partial G), \mathrm{mhd}(\partial G, \partial P)\right), \tag{23}$$

其中，mhd(.)表示相对改进豪斯多夫距离：

$$\mathrm{mhd}(\partial P, \partial G) = \frac{1}{|\partial P|} \sum_{p \in \partial P} \min_{g \in \partial G} (\|p - g\|_2) \tag{24}$$

其中，p 和 g 表示相应离散椭圆上的点（椭圆 ∂P上的点总数为：$|\partial P|$）。

考虑这五项指标，可以相当全面地比较不同椭圆预测方法的性能。

表 I - AKM [19] 数据集上的实验验证。最佳结果以粗体显示

	交并比 (↑)	重叠度 (↑)	Dice (↑)	相对体积差 (↓)	平均豪斯多夫距离 (↓)
Rubino 等人 [18]	0.78±0.06	0.93±0.04	0.88±0.04	0.12±0.08	6.88±5.98
Zins等人 [10]	0.91±0.12	0.96±0.07	0.95±0.08	0.02±0.04	3.93±5.63
我们的方法 ($\mathcal{L}_{\mathrm{W}}$)	0.89±0.05	**0.98±0.02**	0.94±0.03	0.08±0.06	3.48±1.80
我们的方法 ($\mathcal{L}_{\mathrm{W}} + \mathcal{L}_{\mathrm{JS}}$)	**0.93±0.03**	0.97±0.02	**0.96±0.01**	**0.01±0.01**	**2.16±0.66**
本文方法 ($\mathcal{L}_{\mathrm{JS}}$)	**0.93±0.03**	0.97±0.02	**0.96±0.02**	**0.01±0.01**	2.26±0.77

表 II - SPEED+ [13] 数据集上的实验验证。最佳结果以粗体显示

	交并比 (↑)	重叠度 (↑)	Dice (↑)	相对体积差 (↓)	平均豪斯多夫距离 (↓)
Rubino 等人 [18]	0.70±0.13	0.91±0.10	0.82±0.13	0.17±0.17	88.35±119.49
Zins等人 [10]	0.85±0.18	**0.95±0.12**	0.91±0.15	0.07±0.15	54.38±116.38
我们的方法 ($\mathcal{L}_{\mathrm{W}}$)	0.82±0.09	**0.95±0.05**	0.90±0.06	0.08±0.13	9.59±7.18
我们的方法 ($\mathcal{L}_{\mathrm{W}} + \mathcal{L}_{\mathrm{JS}}$)	0.86±0.08	**0.95±0.05**	0.92±0.05	0.03±0.06	7.85±6.86
我们的方法 ($\mathcal{L}_{\mathrm{JS}}$)	**0.87±0.07**	**0.95±0.04**	**0.93±0.05**	**0.01±0.05**	**7.58±6.89**

基线方法 对于由椭球建模的三维物体，我们的方法旨在回归其在图像中的椭圆投影。我们将本方法与另外两种具有代表性的三维感知椭圆回归方法进行了定性和定量比较。值得注意的是，与此相反，大多数二维椭圆回归方法 [3], [4], [5], [6], [7], [2], [8] 旨在检测图像中自然存在的二维椭圆形状。第一种基线方法来自 Zins 等人 [10], ，这是一种两阶段方法：先使用目标检测器（官方实现中采用 Faster R-CNN [6]和 ResNet50 [25] 骨干网络）定位物体，再从裁剪后的图像中回归椭圆几何参数。该方法是其先前工作的改进版本 [9]。第二种方法常用于基于椭球的姿态估计问题（例如， [18]），即在检测到的边界框内拟合一个与坐标轴对齐的椭圆（实验中使用相同的检测模型）。

目标定位结果表 I 和表 II 对比了仅使用 Wasserstein 损失、仅使用 Jensen-Shannon 损失以及结合两者（$\lambda = 1$）优化的模型与 Rubino等人 [18] 和 Zins 等人 [10]提出的方法。结果表明，尽管我们的方法采用了更轻量的骨干网络（ResNet32），但在 AKM 和 SPEED+数据集的五项指标上均优于其他方法。对于 AKM 数据集上的模型，利用 Jensen-Shannon 散度的两种损失都取得了相同的准确率，这表明最重要的优化因素是隐式函数正则化。在 SPEED+上，仅使用 Jensen-Shannon 散度训练的模型表现最佳，因为高斯隐式函数在图像边界之外略有截断，从而导致参数提取出现偏差。

此外，我们的单阶段方法能够以完全可微的方式同时执行定位和参数提取，这是其优势所在。此外，参数提取由新颖的 E-DSNT 不可训练层完成，因此模型更加轻量。最

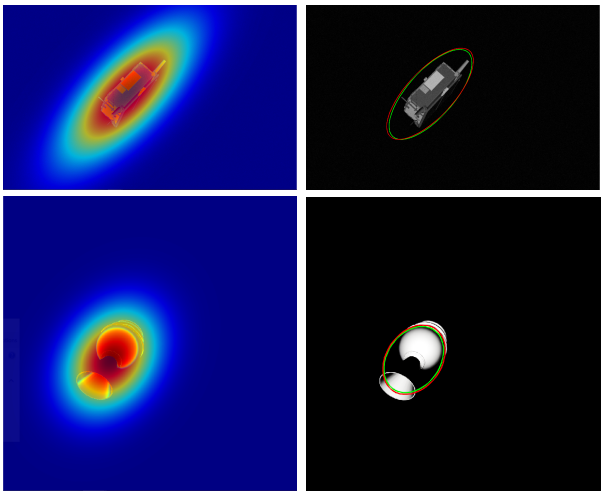


图4. SPEED+ （上）和 AKM（下）数据集上的定性结果。由回归参数得到的隐式占用函数以热力图形式叠加显示（左）；提取出的椭圆（红色）和真实值椭圆（绿色）（右）。

后，它避免了回归不连续角参数（椭圆方向），这与 [9], [10] 及所有其他椭圆检测方法不同。图4所示的定性结果展示了隐式占据函数值、真实值和预测椭圆。

三维重建 为了评估预测结果的三维感知能力，我们将回归椭圆的椭圆参数作为输入，采用一种基于二维椭圆三角测量的三维椭球重建方法 [18]。我们从 SPEED+ 数据集中随机选取了100张图像，并根据不同方法回归得到的椭圆重建椭球。图5显示，与其他方法（蓝色、洋红色）得到的结果相比，根据我们的预测结果重建的椭球（红色）更接近真实椭球（绿色）。此外，我们对200组由50张随机图像构成的子集进行了评估，结果见表III。结果表明，我们的方法能够实现最准确的重建，证明了

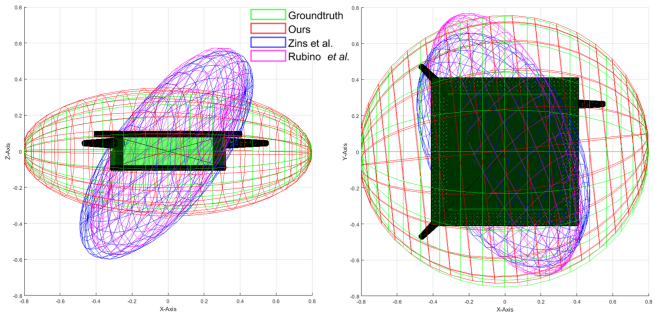


图 5. TANGO 航天器（SPEED+ 数据集）的示意图，其中真实值以绿色表示，我们的方法重建的椭球体以红色表示，Zins 等人 [9] 以蓝色表示，Rubino 等人 [18] 以洋红色表示。重建方法来自 [18]。根据我们的预测结果重建的椭球比其他方法得到的椭球更接近真实椭球。

表 III - SPEED+ [13]数据集上使用椭圆三角测量法 [18]得到的三维重建误差。最佳结果以粗体显示。

	位置	方向	Size
Rubino 等人。 [18]	5.9 厘米	10.2°	11.0cm
Zins 等人 [10]	2.1cm	4.1°	8.0cm
我们的方法 ($\mathcal{L}_{\text{JS}}$)	**1.0厘米**	**0.1°**	**2.4厘米**

回归椭圆的三维感知能力。这对于后续可能进行的六自由度（6DoF）物体姿态估计任务 [26], [27]尤为重要。

V. 结论

本文提出了一种基于高斯隐式占据函数和椭圆标签的完全可微三维感知目标定位方法。我们介绍了如何轻松生成一致的高斯占据标签，以扩展现有姿态数据集，而无需物体的任何 CAD模型。我们还发布了三个公开航天器姿态估计数据集的标签。未来的工作将重点把该二维定位模型集成到端到端的六自由度航天器姿态估计流程中，并在真实数据集上进行评估 [13], [28]。

参考文献

[1] L. Liu、W. Ouyang、X. Wang、P. Fieguth、J. Chen、X. Liu 和 M. Pietik¨ainen，“ 通用目标检测的深度学习：综述 ”，国际计算机视觉期刊，第 128 卷，第 2 期，第 261–318 页，2020 年。[2] W. Dong、P. Roy、C. Peng 和 V. Isler，“Ellipse R-CNN：通过聚类和遮挡学习推断椭圆形物体”，*IEEE* 图像处理汇刊，第 30 卷，第 2193–2206 页，2021 年。[3] Z. Wang、N. Dong、S. D. Rosario、M. Xu、P. Xie 和 E. P. Xing，“眼底图像中视盘与视杯边界的椭圆检测”，载于第 *16* 届 *IEEE* 生物医学成像国际研讨会（*ISBI*），2019 年，第 601–604 页。[4] H. Lin、Z. Li、M. Shih、Y. Sun 和 T. Shen，“ 利用锚框椭圆回归进行眼科诊断中的瞳孔定位”，载于第 *16* 届机器视觉应用国际会议（*MVA*），2019 年，第 1–5 页。[5] Y. Li，“利用高斯提议网络检测病灶边界椭圆”，载于*MICCAI 2019* 研讨会论文集。Springer，2019 年，第 337–344 页。[6] S. Ren、K. He、R. B. Girshick 和 J. Sun，“Faster R-CNN：借助区域提议网络实现实时目标检测”，载于*NIPS*，2015 年，第 91–99 页。

[7] S. Pan、S. Fan、S. W. K. Wong、J. V. Zidek 和 H. Rhodin，“椭圆检测与定位及其在锯材图像木节检测中的应用”，载于 IEEE 冬季计算机视觉应用会议（WACV），2021 年，第 3891–3900 页。[8] T. Wang、C. Lu、M. Shao、X. Yuan 和 S. Xia，“ElDet：一种无锚框通用椭圆目标检测器”，载于亚洲计算机视觉会议（ACCV）论文集，2022 年，第 2580–2595 页。[9] M. Zins、G. Simon 和 M.-O. Berger，“用于基于物体的相机姿态估计的三维感知椭圆预测”，载于第 8 届三维视觉国际会议（3DV 2020）。IEEE，2020 年，第 281–290 页。[10] ——，“基于椭球模型和三维感知椭圆预测的基于物体的视觉相机姿态估计”，《国际计算机视觉期刊》，第 130 卷，第 4 期，第 1107–1126 页，2022 年。[11] ——，“基于水平集的相机姿态估计：利用多组二维/三维椭圆-椭球对应关系”，载于 IEEE/RSJ 智能机器人与系统国际会议（IROS），2022 年。[12] P. F. Proenc¸a 和 Y. Gao，“利用照片级真实感渲染进行航天器姿态估计的深度学习”，载于 IEEE 机器人与自动化国际会议（ICRA），2020 年，第 6007–6013 页。[13] T. H. Park、M. M¨artens、G. Lecuyer、D. Izzo 和 S. D’Amico，“SPEED+：用于跨域差异航天器姿态估计的新一代数据集”，载于 IEEE 航空航天会议（AERO），2022 年，第 1–15 页。[14] R. I. Hartley 和 A. Zisserman，《计算机视觉中的多视图几何》，第 2 版。剑桥大学出版社，2004 年。[15] T. Fan、G. Wang、Y. Li 和 H. Wang，“ MA-Net：用于肝脏和肿瘤分割的多尺度注意力网络”，《IEEE Access》，第 8 卷，第 179 656–179 665 页，2020 年。[16] K. He、X. Zhang、S. Ren 和 J. Sun，“用于图像识别的深度残差学习”，载于 IEEE 计算机视觉与模式识别会议（CVPR）论文集，2016 年 6 月。[17] A. Nibali、Z. He、S. Morgan 和 L. Prendergast，“利用卷积神经网络进行数值坐标回归”，arXiv 预印本 arXiv:1801.07372，2018 年。[18] C. Rubino、M. Crocco 和 A. Del Bue，“基于多视图图像检测结果的三维物体定位”，《IEEE 模式分析与机器智能汇刊》，第 40 卷，第 6 期，第 1281–1294 页，2018 年。[19] A. Rathinam、V. Gaudilliere、L. Pauly 和 D. Aouada，“AKM数据集：无纹理空间目标数据集”，https://zenodo.org/record/7744505，2022 年 9 月。[20] M. Kisantal、S. Sharma、T. H. Park、D. Izzo、M. M¨artens 和 S. D’Amico，“卫星姿态估计挑战赛：数据集、竞赛设计与结果”，《IEEE 航空航天与电子系统汇刊》，第 56 卷，第 5 期，第 4083–4098 页，2020 年。[21] A. Rathinam、V. Gaudilli`ere、L. Pauly 和 D. Aouada，“利用卷积神经网络对已知无纹理空间目标进行姿态估计”，载于第 73 届国际宇航大会（IAC），2022 年。[22] L. Pauly、W. Rharbaoui、C. Shneider、A. Rathinam、V. Gaudilli`ere 和 D. Aouada，“基于深度学习的单目航天器姿态估计综述：现状、局限与前景”，ArXiv 预印本，第 abs/2305.07348 卷，2023 年。 [在线]。可通过以下网址获取：https://doi.org/ 10 48550 2305 07348 。/ arXiv. 。S. M. Catalogue，[23] “Prisma（原型）”，https://www.eoportal.org/satellite-missions/prisma-prototype#target-spacecraft。[24] M.-P. Dubuisson 和 A. Jain，“用于物体匹配的改进豪斯多夫距离”，载于第 12 届模式识别国际会议论文集，第 1 卷，1994 年，第 566–568 页。[25] K. He、X. Zhang、S. Ren 和 J. Sun，“用于图像识别的深度残差学习”，载于 2016 IEEE 计算机视觉与模式识别会议（CVPR），2016 年，第 770–778 页。[26] A. Garcia、M. A. Musallam、V. Gaudilli`ere、E. Ghorbel、K. Al Ismaeil、M. Perez 和 D. Aouada，“LSPNet：面向二维定位的航天器姿态估计神经网络”，载于 IEEE/CVF 计算机视觉与模式识别会议论文集，2021 年，第 2048–2056 页。[27] V. Gaudilli`ere、G. Simon 和 M.-O. Berger，“ Perspective-1-ellipsoid：基于单组椭圆-椭球对应关系的相机姿态估计问题的表述、分析与求解”，《国际计算机视觉期刊》，第 131 卷，第 9 期，第 2446–2470 页，2023 年。[28] A. Rathinam、V. Gaudilli`ere、M. A. Mohamed Ali、M. Ortiz Del Castillo、L. Pauly 和 D. Aouada，“ SPARK 2022数据集：航天器检测与轨迹估计”，2022 年 6 月。 [在线]。可通过以下网址获取：https://doi.org/10.5281/zenodo.6599762