6. Supplementary Experiments

6.1. Keypoint Confidence Regions

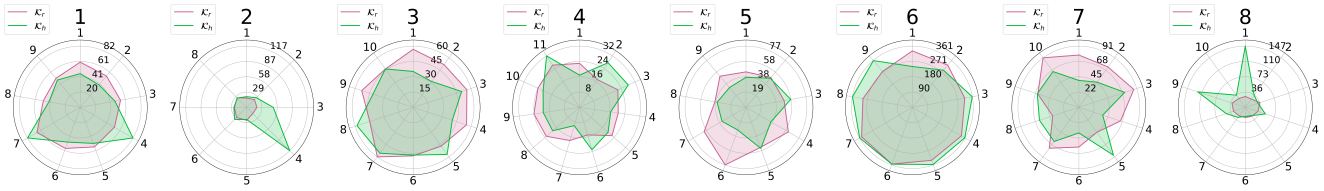

Figure 6. Radar charts illustrating the mean radii of keypoint confidence regions ($\mathcal{K}^h$ and $\mathcal{K}^r$). The eight radar charts correspond to the eight objects in the LMO dataset (from left to right). Each axis in a given radar chart represents one of the specific object's keypoints.

LMO [23] Objects	Ours	[9]		[11]
		$\epsilon = 0.1$	$\epsilon = 0.4$	
ape (1)	76.78	77.70	**79.52**	69.14
can (2)	**91.96**	73.41	75.97	86.09
cat (3)	90.11	87.36	**90.59**	65.12
driller (4)	**89.29**	79.32	83.08	61.44
duck (5)	**84.39**	82.71	82.54	73.06
eggbox (6)	6.85	0	0	**8.43**
glue (7)	69.83	56.49	**71.08**	55.37
holepuncher (8)	**86.44**	81.65	82.89	69.84
mean	**74.45**	67.33	70.71	61.06
SPEED [22]	**97.09**	57.80	57.40	57.46

Table 5. Acc of baseline methods and our approach

LMO [23] Objects	Ours $\mathcal{K}^r$		[9] $\mathcal{K}^h$	
	$\epsilon = 0.1$	$\epsilon = 0.4$	$\epsilon = 0.1$	$\epsilon = 0.4$
1	**90.37**	61.51	88.35	64.87
2	89.97	59.90	**93.54**	61.81
3	91.63	64.63	**92.30**	63.69
4	**91.84**	59.30	90.86	64.74
5	**90.31**	64.09	90.13	64.94
6	88.40	59.90	**88.86**	60.37
7	**92.02**	54.34	91.97	59.83
8	90.66	59.42	**93.31**	66.86
mean	90.65	60.38	**91.16**	63.38
SPEED [22]	**89.66**	61.25	88.88	62.64

Table 6. η^{kpt} of baseline methods and our approach

6.2. Qualitative Pose Confidence Regions

Fig. 7 and Fig. 8 provides SPEED satellites' $\mathcal{K}^r$, $\mathcal{R}_r^d$, and $\mathcal{T}_r^d$ with $\epsilon = 0.1$ and $\epsilon = 0.4$ across varying viewpoints and diverse backgrounds.

Fig. 9 and Fig. 10 provides 8 LMO objects' $\mathcal{P}_\Theta(\mathbf{x}_n|\mathbf{I})$, $\mathcal{K}^r$, $\mathcal{R}_r^d$, and $\mathcal{T}_r^d$ with $\epsilon = 0.1$ and $\epsilon = 0.4$.

6.3. Single-shot PnP

Unlike sampling methods, in this paper we follows [45] utilizing the single-shot PnP weighted by $\boldsymbol{\sigma}_n$, thereby significantly enhancing the algorithm's efficiency. To incorporate the uncertainty into our single-shot PnP framework, we assign weights to $\tilde{\mathbf{x}}_n$ based on $\boldsymbol{\sigma}^n$. The optimization problem for pose estimation can be expressed as follows:

$$\min_{\mathbf{R},\mathbf{t}} \sum_{n=1}^{N} \rho\left(\mathbf{r}(n)^\top \left(\boldsymbol{\sigma}^n\right)^{-1} \mathbf{r}(n)\right), \tag{19}$$

where $\mathbf{r}(i)$ denotes the reprojection error, defined as:

$$\mathbf{r}(n) = \lambda_i \begin{bmatrix} \mathbf{x}^n \\ 1 \end{bmatrix} - \mathbf{K}\left[\mathbf{R}|\mathbf{t}\right] \begin{bmatrix} \mathbf{z}^n \\ 1 \end{bmatrix}. \tag{20}$$

In Eq. (20), λ_i represents the depth value associated with the 3D point $\mathbf{z}^n$, used to scale the projection onto the 2D image plane accurately. In Eq. (19), $\rho(\cdot)$ represents a robust loss function. Consistent with the methodology of Wang *et al.* [45], we employ the Huber loss as the robust estimation function, given by:

$$H(e) = \begin{cases} \frac{1}{2}e^2, & \text{if } e \leq \delta \\ \delta(|e| - \frac{1}{2}\delta), & \text{otherwise} \end{cases}. \tag{21}$$

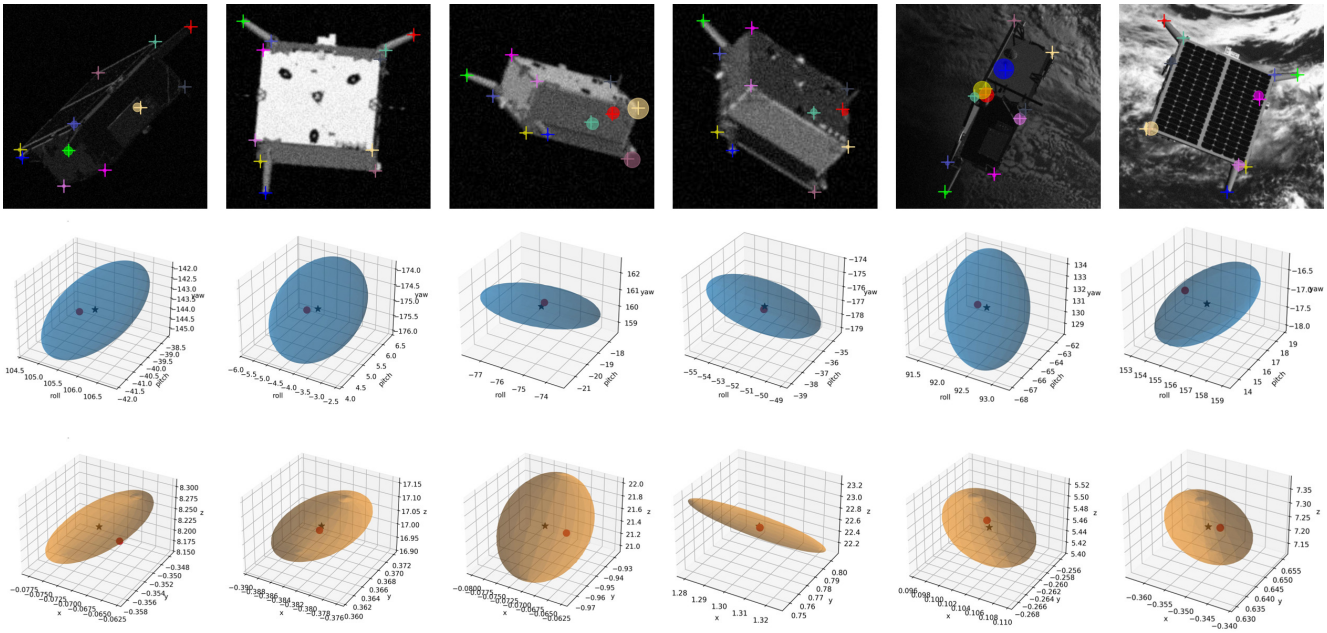

Figure 7. $\mathcal{K}^r$, $\mathcal{R}_r^d$, and $\mathcal{T}_r^d$ (corresponding to rows 1 to 3) for six images (corresponding to columns 1 to 6) in the SPEED dataset when $\epsilon = 0.1$.

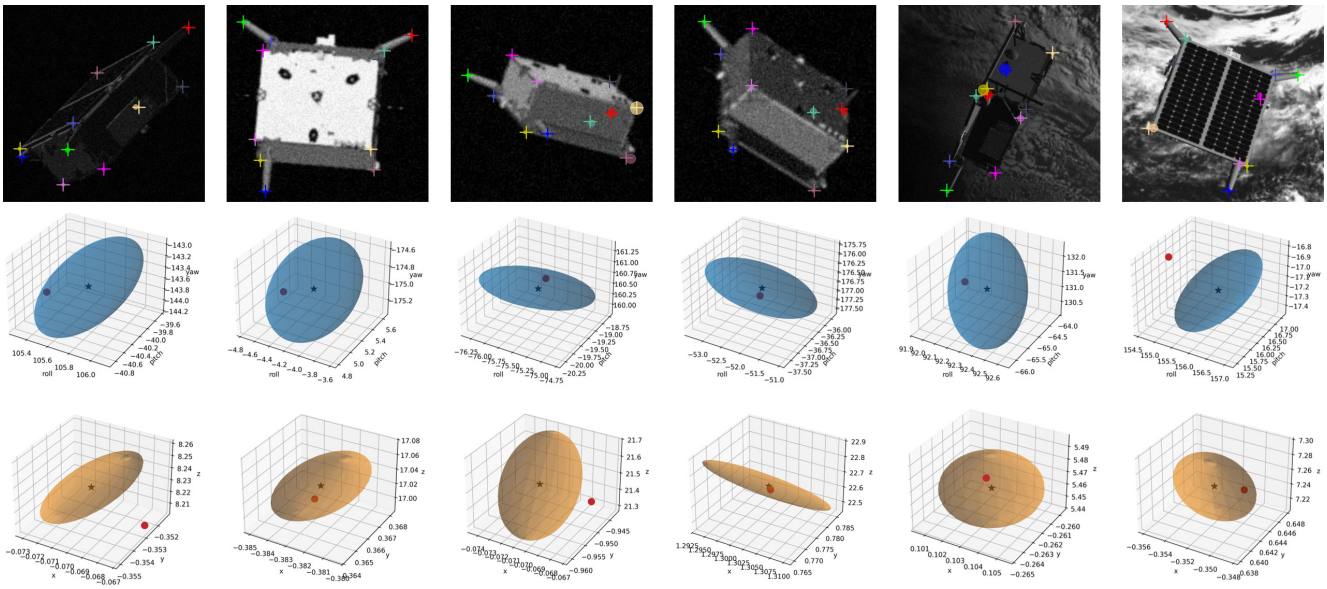

Figure 8. $\mathcal{K}^r$, $\mathcal{R}_r^d$, and $\mathcal{T}_r^d$ (corresponding to rows 1 to 3) for six images (corresponding to columns 1 to 6) in the SPEED dataset when $\epsilon = 0.4$.

Conventional PnP algorithms are prone to inaccuracies from poorly localized keypoints. By integrating uncertainty weights into our single-shot PnP approach, we enhance pose estimation accuracy and overcome the limitations of traditional methods.

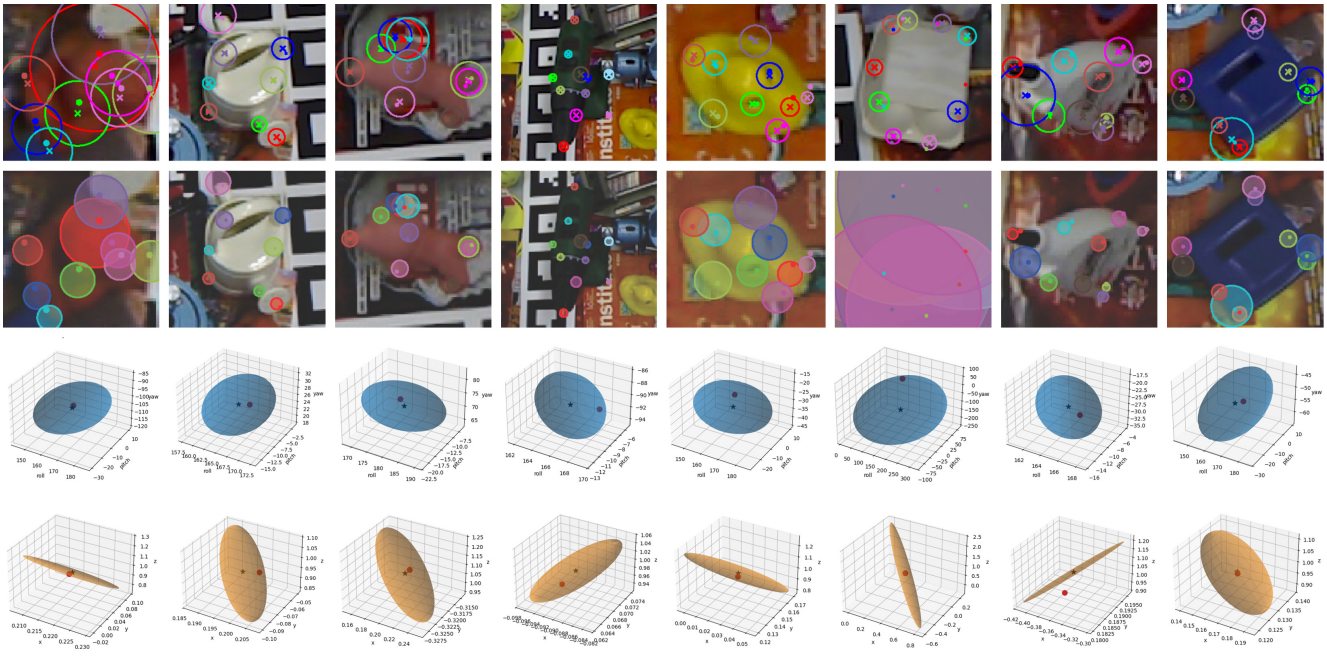

Figure 9. $\mathcal{P}_\Theta(\mathbf{x}_n|\mathbf{I})$, $\mathcal{K}^r$, $\mathcal{R}_r^d$, and $\mathcal{T}_r^d$ (corresponding to rows 1 to 4) for eight objects (corresponding to columns 1 to 8) in some image of the LMO dataset when $\epsilon = 0.1$.

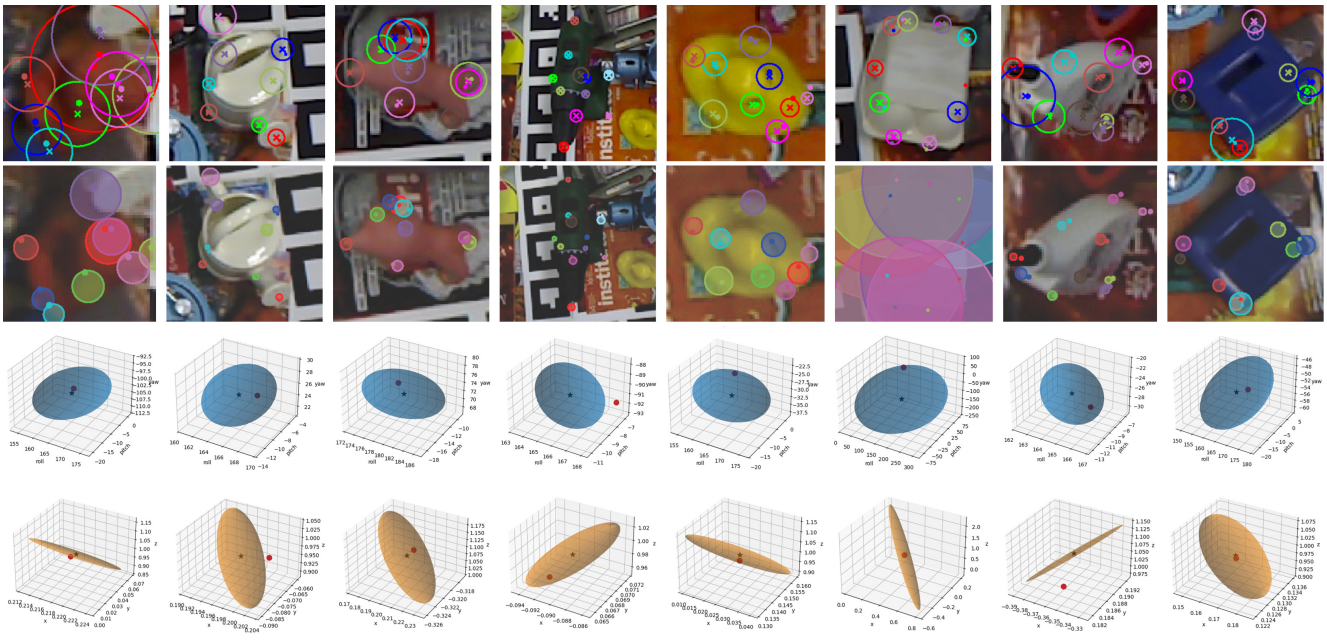

Figure 10. $\mathcal{P}_\Theta(\mathbf{x}_n|\mathbf{I})$, $\mathcal{K}^r$, $\mathcal{R}_r^d$, and $\mathcal{T}_r^d$ (corresponding to rows 1 to 4) for eight objects (corresponding to columns 1 to 8) in some image of the LMO dataset when $\epsilon = 0.4$.