

Supplementary Material

Pose Estimation for Objects with Rotational Symmetry

Enric Corona¹, Kaustav Kundu¹, Sanja Fidler^{2*†}

In the supplementary material, we provide a more detailed discussion about rotational symmetry and provide proofs for our claims in Sec. 1. We further provide additional results of our approach in Sec. 2.

1 Rotational Symmetry

We start by introducing the notation and definitions as per Sec. 1.1, followed by proofs of claims in Sec. 1.2. In Sec. 1.3, we discuss two of the limitations in our approach of reasoning about the rotational symmetry.

1.1 Notation and Definitions

Rotation Matrix. We denote a rotation for an angle ϕ around an axis θ using a matrix $\mathbf{R}_\theta(\phi)$. For example, if the axis of rotation is the X-axis, then

$$\mathbf{R}_X(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi & \cos \phi \end{bmatrix}$$

Order of Rotational Symmetry. We say that an object has an n order of rotational symmetry around the axis θ , *i.e.*, $\mathcal{O}(\theta) = n$, when its 3D shape is equivalent to its shape rotated by $\mathbf{R}_\theta\left(\frac{2\pi i}{n}\right)$, $\forall i \in \{0, \dots, n-1\}$.

The minimum value of $\mathcal{O}(\theta)$ is 1, and attained for objects non-symmetric around axis θ . The maximum value is ∞ , which indicates that the 3D shape is equivalent when rotated by any angle around its axis of symmetry. This symmetry is also referred to as the revolution symmetry [1]. In Fig. 1, we can see an example of our rotational order definition. For a 3D model shown in Fig. 1 (a), the rotational order about the Y axis is 2, *i.e.*, $\mathcal{O}(\mathbf{Y}) = 2$. Thus for any viewpoint v (cyan) in Fig. 1 (b), if we rotate it by π about the Y-axis to form, $v_\pi = \mathbf{R}_Y(\pi)v$, the 3D shapes will be equivalent

^{*1}Enric Corona and Kaustav Kundu are with Department of Computer Science, University of Toronto ecorona@cs.toronto.edu

^{†2}Sanja Fidler is with Department of Computer Science, University of Toronto, and the Vector Institute fidler@cs.toronto.edu

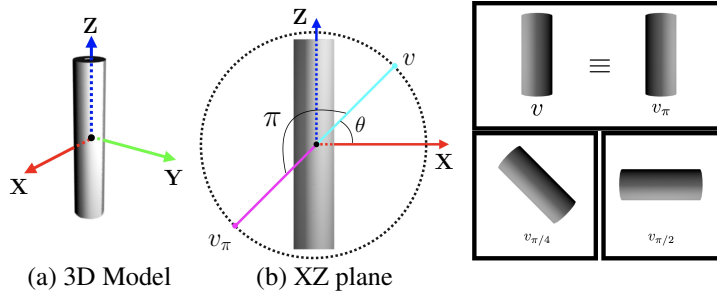


Figure 1: Order of Rotational Symmetry

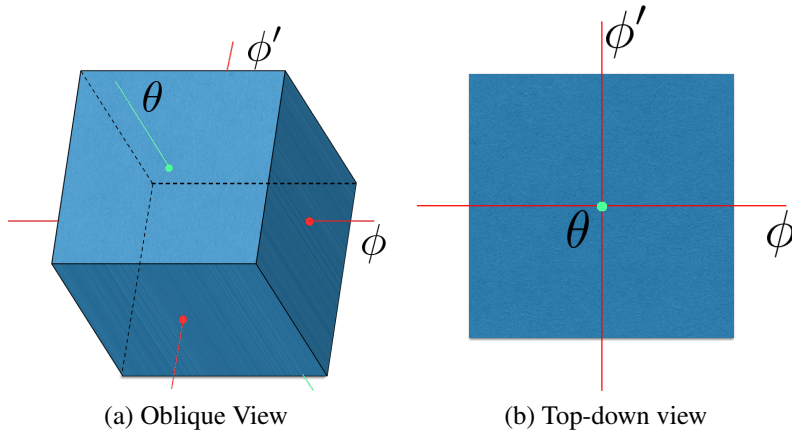


Figure 2: Illustration for Claim 1. For the two axes shown, $\mathcal{O}_{cube}(\theta) = 4$ and $\mathcal{O}_{cube}(\phi) = 4$. Equivalent views repeat every $\frac{2\pi i}{n_\theta}$ when rotating around the axis θ . From Claim 1, an axis $\phi' = \mathbf{R}_\theta(\frac{\pi}{2})\phi$ (for $i = 1$) will have rotational order, $\mathcal{O}_{cube}(\phi') = 4$.

(Fig. 1 (right)). The 3D shape in any other viewpoint (such as, $v_{\pi/4}$ or $v_{\pi/2}$) will not be equivalent to that of v . Similarly, we have $\mathcal{O}(\mathbf{Z}) = \infty$. In our paper, we only consider the values of rotational order to be one of $\{1, 2, 4, \infty\}$, however, our method will not depend on this choice.

Equivalent Viewpoint Sets. Let us define the set of all pairs of equivalent viewpoints as $E_o(\mathbf{Y}) = \{(i, j) | v_j = \mathcal{R}_\theta(\pi)v_i\}$, with an symmetry order $o \in \{2, 3, \infty\}$. Note that $E_1(\theta)$ is a null set (object is asymmetric). In our case, we have $E_2(\theta) \subset E_4(\theta) \subset E_\infty(\theta)$ and $E_3(\theta) \subset E_\infty(\theta)$.

1.2 Geometrical Constraints on Orders of Rotational Symmetry about Multiple Axes

Claim 1. For an object x let two non-collinear axes θ and ϕ have orders of rotational symmetry, $\mathcal{O}_x(\theta) = n_\theta$ and $\mathcal{O}_x(\phi) = n_\phi$. Then, $\forall i \in \{0, \dots, n_\theta\}$, the order of symmetry around the axis $\mathbf{R}_\theta\left(\frac{2\pi i}{n_\theta}\right)\phi$ is also n_ϕ , i.e., $\mathcal{O}_x\left(\mathbf{R}_\theta\left(\frac{2\pi i}{n_\theta}\right)\phi\right) = n_\phi$.

Similarly $\forall j \in \{0, \dots, n_\phi\}$, $\mathcal{O}_x\left(\mathbf{R}_\theta\left(\frac{2\pi j}{n_\phi}\right)\theta\right) = n_\theta$. We provide an illustration of this statement with an example in Fig. 2.

Proof. Since $\mathcal{O}_x(\phi) = n_\phi$, for any viewpoint v , the 3D shape of the object x , is equivalent to that from the viewing direction, $v_\phi^j = \mathbf{R}_\phi\left(\frac{2\pi j}{n_\phi}\right)v$, $\forall j \in \{1, \dots, n_\phi - 1\}$. Also $\mathcal{O}_x(\theta) = n_\theta$ implies that x is equivalent from the viewpoints, v and $v_\theta^i = \mathbf{R}_\theta\left(\frac{2\pi i}{n_\theta}\right)v$ for any $i \in \{1, \dots, n_\theta\}$. Similarly, x is equivalent from v_ϕ^j and $v_{\theta,\phi}^{i,j} = \mathbf{R}_\theta\left(\frac{2\pi i}{n_\theta}\right)v_\phi^j$. If we can prove this claim for any arbitrary, i and j , then we are done.

For ease of notation, we replace the symbols, $v_\phi^j, v_\theta^i, v_{\theta,\phi}^{i,j}$ by $v_\phi, v_\theta, v_{\theta,\phi}$, respectively.

Now, let us consider a point p_ϕ along the ϕ axis, an unit distance away from the origin, O . We rotate the polyhedron, P (from the vertices, v, v_ϕ, p_ϕ and O) by an angle of $\frac{2\pi i}{n_\theta}$ around the axis, θ , i.e., each of the vertices are rotated by $\mathbf{R}_\theta\left(\frac{2\pi i}{n_\theta}\right)$. After the rotation, v and v_ϕ would coincide with v_θ and $v_{\theta,\phi}$ respectively. The axis, ϕ would become $\phi' = \mathbf{R}_\theta\left(\frac{2\pi i}{n_\theta}\right)\phi$ and p_ϕ would be transformed to $p_{\phi'} = \mathbf{R}_\theta\left(\frac{2\pi i}{n_\theta}\right)p_\phi$.

Since rotation is an isometry, the angles are preserved under the rotation transformation. Thus, the angle between v_θ and $v_{\theta,\phi}$ would remain as $\frac{2\pi j}{n_\phi}$, i.e., $\angle v_\theta O v_{\theta,\phi} = \frac{2\pi j}{n_\phi}$. Similarly, $\angle v_\theta O p_{\phi'} = \angle v_{\theta,\phi} O p_{\phi'} = \pi/2$. Thus, rotating v_θ by $\frac{2\pi j}{n_\phi}$ around the axis, ϕ' would give us $v_{\theta,\phi}$.

Now since, the 3D shape is equivalent between the following pair of viewpoints, (v, v_ϕ) , (v, v_θ) and $(v_\phi, v_{\theta,\phi})$, thus it would be equivalent from v_θ and $v_{\theta,\phi}$. Thus if $\mathcal{O}_x(\phi) = n_\phi$, then $\mathcal{O}_x(\phi') = n_\phi$.

Similarly, we can prove that, $\forall j \in \{0, \dots, n_\phi\}$, $\mathcal{O}_x\left(\mathbf{R}_\theta\left(\frac{2\pi j}{n_\phi}\right)\theta\right) = n_\theta$. \square

Please note that here the equivalence corresponds to that of the entire 3D shape and not the projection of the 3d shape onto a plane at a particular viewpoint. Considering the latter is more difficult, we leave it for future work.

Corollary 1. An object is a sphere iff two non-collinear axes have infinite order rotational symmetry.

Proof. If an object x is a sphere, it is symmetric about all directions. Thus it has infinite order rotational symmetry in any pair of axes. Let us now prove the opposite, i.e., if

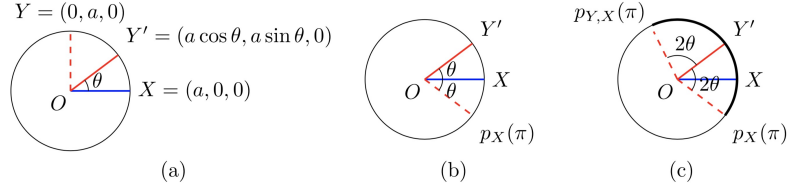


Figure 3: Illustration for Corollary 1. Please see text for details.

two non-collinear axes have infinite order rotational symmetry, then the object has to be a sphere.

Case 1: Let us first prove this corollary for the special case where the two axes are orthogonal to each other. That is, we aim to show that if any two orthogonal axes have infinite order rotational symmetry, then the object is a sphere. Let \mathbf{X} and \mathbf{Z} axes have infinite orders of rotational symmetry, *i.e.*, $\mathcal{O}_x(\mathbf{X}) = \mathcal{O}_x(\mathbf{Z}) = \infty$. Thus from Claim 1, $\mathcal{O}_x\left(\mathcal{R}_{\mathbf{Z}}\left(\frac{\pi}{2}\right)\mathbf{X}\right) = \mathcal{O}_x(\mathbf{Y}) = \infty$. Since all orthogonal axes have infinite order of rotational symmetry, then the object's shape is equivalent when viewed from any direction. This is only possible for a sphere.

Case 2: Here we prove the corollary for the more general case where the two non-collinear axes are not orthogonal to each other. If we can show that there exist two orthogonal axes which have infinite order of rotational symmetry, then we can follow Case 1 and claim that the object is a sphere.

We let the two axes \mathbf{X} and \mathbf{Y}' lie on the $\mathbf{X} - \mathbf{Y}$ plane with an angle of θ between them and $\mathcal{O}_x(\mathbf{X}) = \mathcal{O}_x(\mathbf{Y}) = \infty$. Let, $X = (a, 0, 0)$ and $Y' = (a \cos \theta, a \sin \theta, 0)$, as shown in Fig. 3 (a). The rotation of Y' about X by an angle α has the following form:

$$p_X(\alpha) = \mathcal{R}_{\mathbf{X}}(\alpha)Y' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} a \cos \theta \\ a \sin \theta \\ 0 \end{bmatrix} = \begin{bmatrix} a \cos \theta \\ a \sin \theta \cos \alpha \\ a \sin \theta \sin \alpha \end{bmatrix}$$

Here, $p_X(\alpha)$ is the parametric equation of a circle on the $\mathbf{Y} - \mathbf{Z}$ plane with the center at $(a \cos \theta, 0, 0)$ and radius $a \sin \theta$. The line joining Y' and $p_X(\pi) = (a \cos \theta, -a \sin \theta, 0)$ will form the diameter of this circle, as shown in Fig. 3 (b). Thus, $\mathcal{O}_x(p_X(\alpha)) = \infty, \forall \alpha \in [0, 2\pi)$.

Similarly, for every $p_X(\alpha)$ we can rotate it about the \mathbf{Y}' axis and those points will have rotational order as ∞ . Every point in the solid arc between $p_{Y',X}(\pi)$ and $p_X(\pi)$ can be obtained from a point $p_X(\alpha)$ rotated on Y' . Thus all the points along the solid arc (lying on the \mathbf{XY} plane) shown in Fig. 3 (c) will have order of rotational symmetry ∞ . We can continue this process until the solid arc crosses Y (as shown in Fig. 3 (a)) the axis orthogonal to \mathbf{X} (on the $\mathbf{X} - \mathbf{Z}$ plane) and this will happen for any $\theta > 0$ as the arc keeps getting bigger. Thus, $\mathcal{O}(Y) = \infty$. Since two orthogonal axes have rotational orders as ∞ , we can use the previous case to show that this object can only be a sphere. \square

Corollary 2. *If an object x is not a sphere, then the following conditions must hold:*

- (a) *It can have up to one axis with infinite order rotational symmetry*
- (b) *If an axis θ has infinite order rotational symmetry, then the order of symmetry of any axis not orthogonal to θ can only be one.*
- (c) *If an axis θ has infinite order rotational symmetry, then the order of symmetry of any axis orthogonal to θ can be a maximum of two.*

Proof.

- (a) This follows directly from Corollary 1.
- (b) Let us assume that an axis ϕ is not orthogonal to θ and $\mathcal{O}_x(\theta) = \infty, \mathcal{O}_x(\phi) = n > 1$. Also, let us rotate θ to $\theta' = \mathcal{R}_\phi \left(\frac{2\pi j}{n} \right) \theta$ for some, $j \in \{1, \dots, n-1\}$. Then, from Claim 1, $\mathcal{O}_x(\theta') = \infty$. But now two non-collinear axes, θ and θ' have infinite orders of symmetry. From Corollary 1, this cannot be true for a non-spherical object. Thus we have a contradiction. Thus $n = 1$.
- (c) It can be proved similarly to the previous part, except that with $n = 2$, the two axes are collinear. Thus n can be either 1 or 2. \square

Since in our experiments none of the objects is a perfect sphere, we use these constraints to improve the accuracy of our symmetry predicting network.

1.3 Limitations of our approach

To reason about rotational symmetry, we use the notion of equivalence between 3D shapes. However, in most practical settings and in our experiments the input is a 2D image, which is the projection of the 3D shape onto an image plane. Thus the occluded part of the 3D shape is not visible in the input. For the viewpoints from which the 3D shapes are equivalent, their corresponding projections will also be equivalent. However the opposite is not true, *i.e.*, it is not necessary that two equivalent projections will have equivalent 3D shapes from their corresponding viewpoints. This problem arises because the back-projecting a 2D image into 3D can have infinite solutions. This cannot be handled by our derivations here.

Moreover, while considering rotational symmetry, we establish equivalence between pairs of viewpoints which are rotations about either **X**, **Y** or **Z** axes. Since we do not reason about pairs of viewpoints which are rotated about any arbitrary axis, our approach does not avoid all the false negative examples in the training data. A trivial extension of our approach to compute rotational orders about any arbitrary axis would scale linearly in both computation and memory costs, which makes this unfeasible.

2 Additional Results

We provide additional quantitative results in Sec. 2.1 and qualitative results in Sec. 2.2.

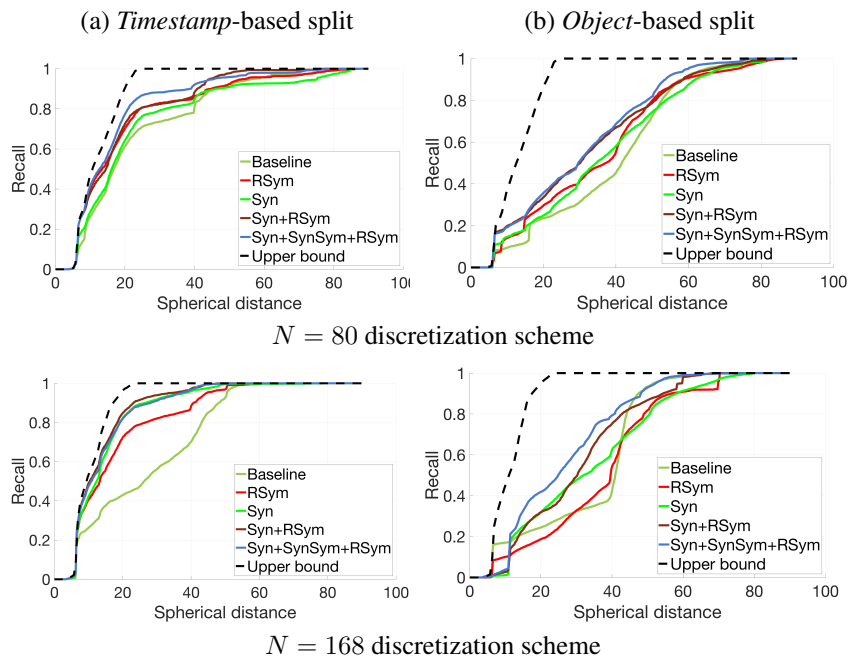


Figure 4: Recall vs spherical distance

2.1 Pose Estimation

In Fig. 4(a) and (b), we plot recall vs the spherical distance between the predicted viewpoint and the GT viewpoint. The first and second rows depict the results from the $N = 80$ and $N = 168$ discretization schemes respectively. We can see that across different discretization schemes, reasoning about rotational symmetry on a large dataset is essential for achieving a good generalization performance.

2.2 Qualitative Results

Rotational Symmetry Prediction. We show examples of the CAD models obtained along with their inferred symmetry in Fig. 5 and Fig. 6. One of the primary reasons for failure is the non-alignment of viewpoints due the discretization. Another reason of failure is that examples of certain order classes are not present in training. For example, the objects in the bottom left of Fig. 5 (xxi, xxii, xxiii) has $\mathcal{O}(\mathbf{Z})$ as 8, 12 and 13 respectively, which was not present in the training set. The orders inferred in all these cases were ∞ . This leads to false positive examples when training for pose estimation.

Since our input is a 2D image, but we reason about the equivalence of 3D shapes, it can confuse the network. An example of this can be seen for the order prediction along the \mathbf{Z} axis in Fig. 5 (ii). A viewpoint from the positive \mathbf{Z} direction would indicate the order to be 4, but from the negative \mathbf{Z} direction it should be inferred as ∞ . However

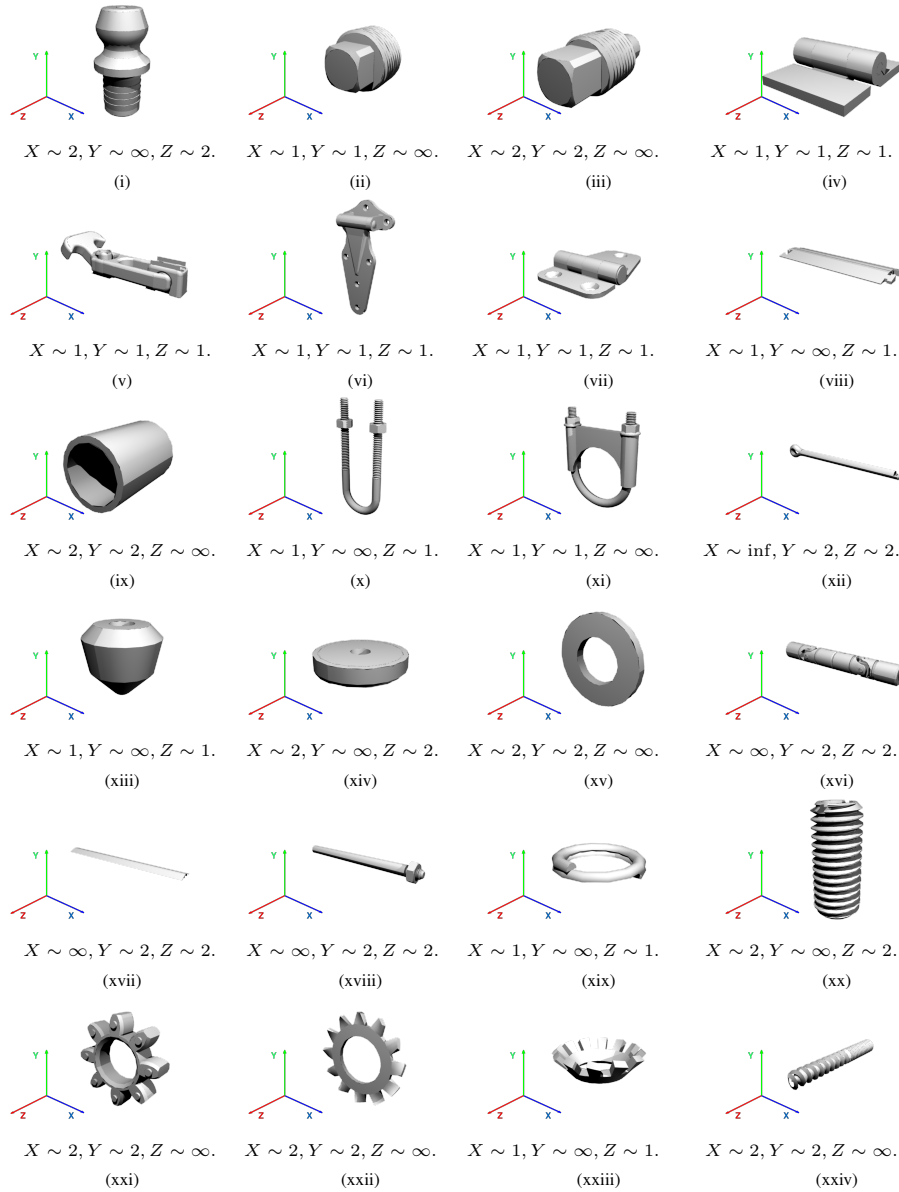


Figure 5: Symmetry Prediction Results

the network in our approach predicts a single prediction of ∞ .

Pose Estimation. Fig. 7 depicts successful results of our approach on the **real dataset**. We also show the successful results on the validation set of the **synthetic dataset** in Fig. 8, Fig. 9 and Fig. 10. For the input image shown in the left column, the depth image of ground truth (GT) coarse pose, followed by top-4 predictions from our ap-



Figure 6: Symmetry Prediction Results

proach. In the last column, we mention the distance of the best match from the top-4 predictions to the ground truth, $d_{\text{rot, best}}^{\text{sym}}$. The results have been sorted in decreasing order of performance of $d_{\text{rot, best}}^{\text{sym}}$. We also draw a green colored box around the best match among the top-4 predictions.

Failure Cases: We show the failure cases in Fig. 11. We draw an orange colored box around the best performing match among the top-4 predictions. In some cases

(rows (a) and (b)) even though none of our top-4 predictions match with the GT, the spherical distance can be very close to that of the GT.

Most of the errors are due to the coarse discretization (rows (d) and (f)). If the actual pose lies in between two neighboring viewpoints, some discriminative parts may not be visible from either of the coarse viewpoints. This can lead to confusion in the matching network. Failure cases in rows (c), (e) and (i) show that the top-k predictions have similar orientation, but the differences in the intricate details is what differentiates it from the GT pose. For synthetic objects, shadows and occlusions can make the problem challenging (rows (g) and (h)).











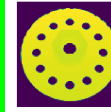
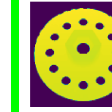




































Input	GT	Top-4 Predictions				$d_{\text{rot, best}}^{\text{sym}}$
						4.62°
						5.93°
						5.94°
						6.37°
						8.21°
						8.40°
						12.68°
						12.78°

Figure 7: **Successful Results:** Qualitative results for pose estimation on the **real** dataset.

























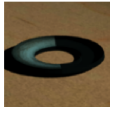



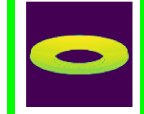















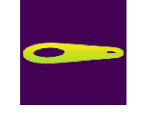

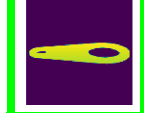
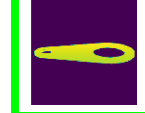
Input	GT	Top-4 Predictions				$d_{rot, best}^{sym}$
						1.49°
						2.02°
						2.34°
						2.35°
						2.79°
						3.08°
						3.34°
						3.37°

Figure 8: **Successful Results:** Qualitative results for pose estimation on the **synthetic** dataset.








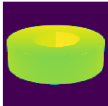
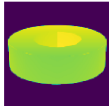
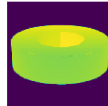
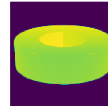
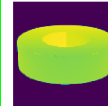








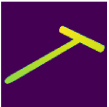









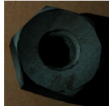
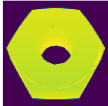
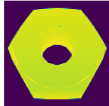
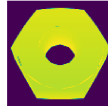
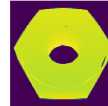
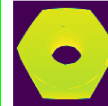

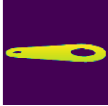
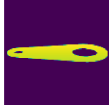
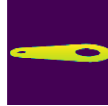








Input	GT	Top-4 Predictions				$d_{\text{rot, best}}^{\text{sym}}$
						4.19°
						4.23°
						4.26°
						4.42°
						4.76°
						5.85°
						6.52°
						6.96°

Figure 9: **Successful Results:** Qualitative results for pose estimation on the **synthetic** dataset.





















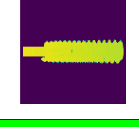
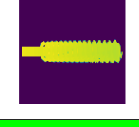
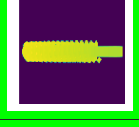
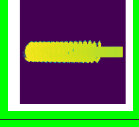


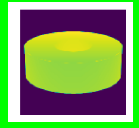
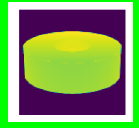
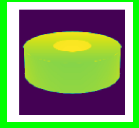
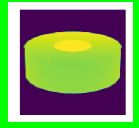
Input	GT	Top-4 Predictions				$d_{rot, best}^{sym}$
						7.73°
						8.36°
						8.63°
						8.64°
						9.21°

Figure 10: **Successful Results:** Qualitative results for pose estimation on the **synthetic** dataset.















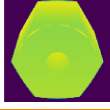



























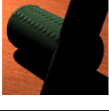



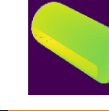


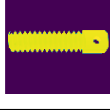
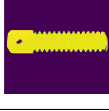
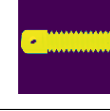
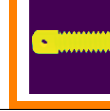
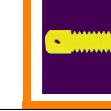
	Input	GT	Top-4 Predictions				$d_{rot, best}^{sym}$
(a)							8.61°
(b)							9.75°
(c)							15.38°
(d)							15.78°
(e)							17.05°
(f)							21.14°
(g)							39.77°
(h)							44.65°
(i)							87.49°

Figure 11: **Failure Cases:** Qualitative results for pose estimation.

References

- [1] R. Brégier, F. Devernay, L. Leyrit, J. L. Crowley, and S.-E. Siléane. Symmetry aware evaluation of 3d object detection and pose estimation in scenes of many parts in bulk. In *CVPR*, pages 2209–2218, 2017. [1](#)