6 Supplementary materials

More qualitative results and details regarding the network architecture are provided in this supplementary.

6.1 Qualitative results of dense correspondences between shapes

We randomly sample some pairs of shapes from different categories and visualize the obtained dense correspondences between each pair of shapes in Figs. 6, 7, and 8. For better visualization, we color the point set by reducing the dimension of the pointwise representations to 3 using t-SNE [27]. Only a sparse set of correspondence links are sampled and visualized. Some of the undesired correspondence links are highlighted by overlaying red blocks on them. These undesired correspondences are mainly due to: 1) highly similar local structures (see the airplane example where the wing tip and the stablizer tip at tail are linked); 2) inconsistent topological structures (see the two examples in chairs).

6.2 Network architecture

Point-based network. The PointNet++ architecture for our pre-trained network consists of three set abstraction (SA) layers, three feature propagation (FP) layers, and a fully-connected layer at the end. The multi-scale grouping strategy is used in the adopted network. An input point cloud with 1024 points is fed to the PointNet++ backbone. In the experiments of Keypoint transfer and Supervised part segmentation, point normals are used as input to the network. The PointNet++ then processes the point cloud to produce a pointwise 128-dimensional features as output. Finally, the FC layer reduces the 128-dimensional output from the PointNet++ backbone to a 64-dimensional output, yielding the final pointwise representaitons. The details are listed in Tab. 5:

Table 5. Network architecture

Layer	Channel sizes of MLPs	# Input points	Multi-scale radii
SA1	$[32, 32, 64], [64, 64, 128], [64, 96, 128]$	512	$[0.1, 0.2, 0.4]$
SA2	$[128, 128, 256], [128, 196, 256]$	256	$[0.4, 0.8]$
SA3	$[256, 512, 1024]$	n.a.	n.a.
FP3	$[512, 256]$	n.a.	n.a.
FP2	$[256, 256]$	n.a.	n.a.
FP1	$[128, 128]$	n.a.	n.a.
FC	$[128, 64]$	n.a.	n.a.

Self-attention module. The set attention blocks proposed in [25] (Equation 8, [25]) are adopted as a self-attention module and inserted in-between each adjacent layers listed in Tab. 5. We instantiate this module with 4 heads and the LayerNorm operation.

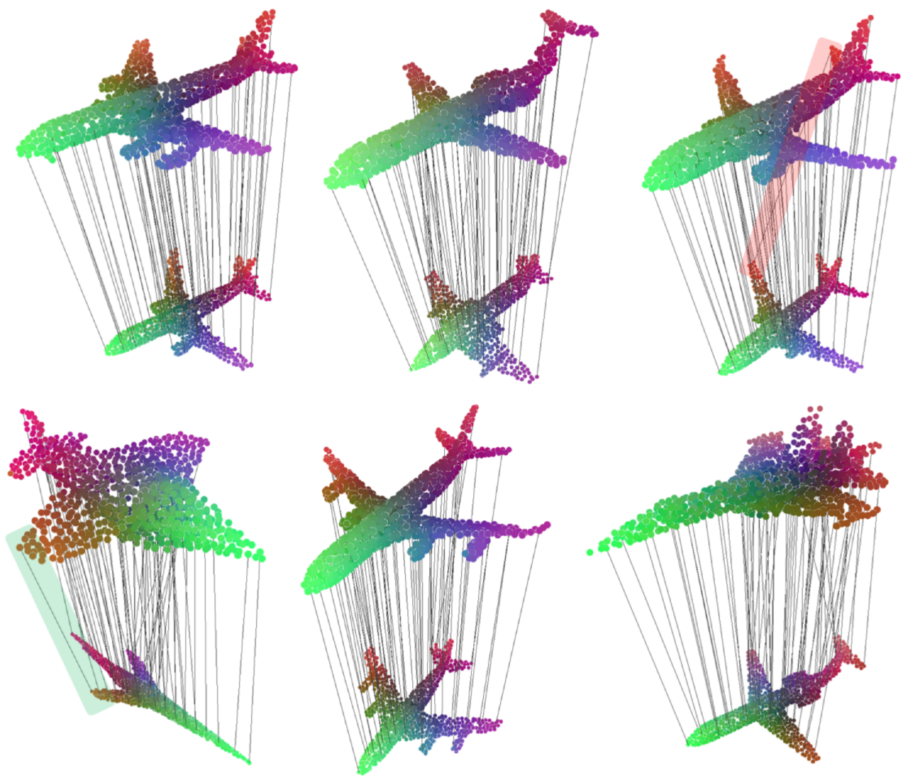

Fig. 6. Dense correspondences between different shapes from the Airplane category. Highly similar local structures (e.g. the wing tip and the stabilizer tip at tail highlighted by the red block in top-right) lead to undesired correspondences. Even the shapes are quite different from each other in bottom left, correct correspondences (highlighted in green) are obtained, showing the robustness of the proposed method in handling inter-instance variation

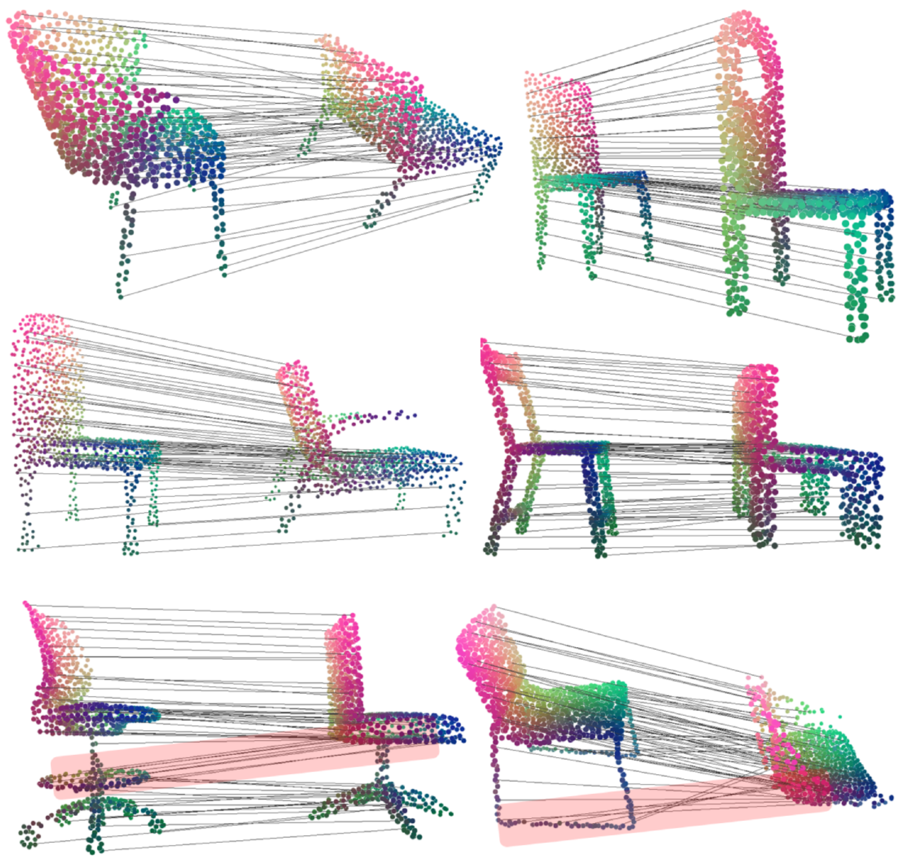

Fig. 7. Dense correspondences between different shapes from the Chair category. Inconsistent topological structures (see the two examples at the bottom row) lead to non-intuitive correspondences.

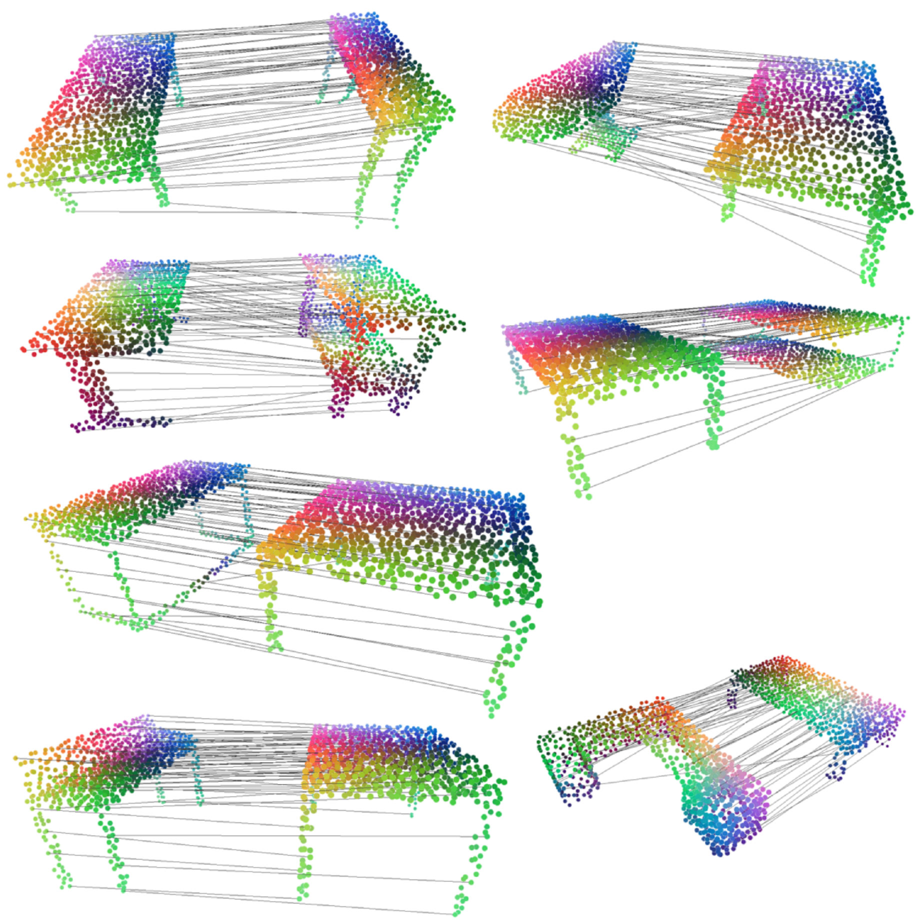

Fig. 8. Dense correspondences between different shapes from the Table category