IPD-Net: SO(3) Invariant Primitive Decompositional Network for 3D Point Clouds

Ramesh Ashok Tabib Nitishkumar Upasi Tejas Anvekar Dikshit Hegde
Uma Mudenagudi
Center of Excellence in Visual Intelligence (CEVI), KLE Technological University,
Hubballi, Karnataka, India
ramesh_t@kletech.ac.in, nitishkumarupasi@gmail.com, anvekartejas@gmail.com,
dikshithegde@gmail.com, uma@kletech.ac.in

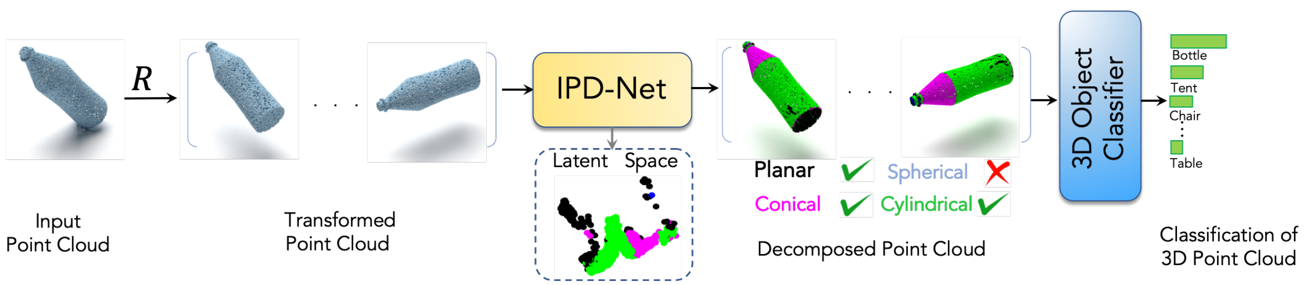

Figure 1. Point cloud data is inherently unstructured and requires a rotation-invariant representation for human-like cognition in machines. Our proposed method, IPD-Net, is a $SO(3)$ invariant framework for decomposition of a point cloud, ensuring robustness to rotations as represented in the latent space. This figure illustrates the need for $SO(3)$ invariant representation and highlights the effectiveness of IPD-Net in achieving rotation-invariant decomposition.

Abstract

In this paper, we propose IPD-Net: Invariant Primitive Decompositional Network, a $SO(3)$ invariant framework for decomposition of a point cloud. The human cognitive system is able to identify and interpret familiar objects regardless of their orientation and abstraction. Recent research aims to bring this capability to machines for understanding the 3D world. In this work, we present a framework inspired by human cognition to decompose point clouds into four primitive 3D shapes (plane, cylinder, cone, and sphere) and enable machines to understand the objects irrespective of its orientations. We employ Implicit Invariant Features (IIF) to learn local geometric relations by implicitly representing the point cloud with enhanced geometric information invariant towards $SO(3)$ rotations. We also use Spatial Rectification Unit (SRU) to extract invariant global signatures. We demonstrate the results of our proposed methodology for $SO(3)$ invariant decomposition on TraceParts Dataset, and show the generalizability of proposed IPD-Net as plugin for downstream task on classification of point clouds. We compare the results of classification with state-of-the-art methods on benchmark dataset (ModelNet40).

1. Introduction

In this paper, we propose IPD-Net: a $SO(3)$ Invariant Framework for understanding the primitive geometry of a 3D point cloud as shown in Figure 1. In recent years, 3D point cloud have begin to play an important role in real world application including SLAM [26] [14], MetaVerse, digitization of heritage sites towards presentation in AR/VR/XR/MR [28] [37] [42] [39], self driving assistance. Towards understanding of point clouds, there is need for an efficient model to analysis the underlying geometry of the point cloud. Large research efforts have been done on solving 3D vision problems with point cloud rather than voxels [52] and multi-view images [41] because of their limitations such as memory footprint. Point clouds are unstructured and unordered in nature making it difficult to analysis the object. Unlike images, which have a defined grid, 3D points clouds are unstructured and unordered making the process difficult through deep learning where we cannot use naive convolution networks. To address this, authors in [33] proposes to use shared-mlp to handle permutation invariance, but each points are processed individually. For understanding the local information, authors in [34] [45] uses nearest neighbours to understand the local information and global information as a hierarchical local features.

The aforementioned methods fail to extract geometrical information from a complex point cloud as these method are designed to extract semantically similar features. Human's cognitive system understand complex objects around us by breaking them down into simple atomic elements or primitives, these primitives can be a higher level interpretable abstraction. Many of downstream task such as point cloud registration [46] and 3D shape retrieval methods [23] often rely on extracting geometric features like keypoints, descriptors, and geometric relationships between points. These features can then be used to match and compare 3D shapes for various applications such as object recognition [16], pose estimation [15], and 3D reconstruction [25].

Exploration towards different primitives like 3D polyhedral shapes [36], generalised cylinders [4], aeons [34], and superquadrics [29] have been done previously where [30], [31] uses cuboidal and superquadratics for 3D shape parsing. Author in [6] uses primitives like squares, circles triangles from from simple hand written drawing are used for graphic programme synthesis. RANSAC [8] and its variants [44] [24] [27] [10] [11] use the christoffel symbols and classical methodology for the decomposition into basic shapes. The main challenge in these method is consideration of prominent features towards decomposition.

To address this, many works have emerged by fitting the point cloud using parametric features and shape fitting [20] [40]. Alternative works decomposes point cloud into basic primitive shapes (plane, cylindrical, cone, and spherical). Each and every geometric shape can be derived from basic primitive shapes. Intuitively extraction of basic primitive features is as good as understanding the morphology of the point cloud, facilitating better representation, generalizability, robustness, scalability, and explainability of point cloud deep neural representation. Towards this, authors in ABD-Net [17] and PointDCCNet [18] propose a methodology in which they split the end-to-end point cloud down stream task into point cloud decomposition followed with a point cloud downstream task. Although the superior performance on decomposition task, these methods are susceptible to rotations of the point cloud. Human cognition is able to interpret and identify familiar objects with any orientation and at any form of abstraction [3] [2]. The idea is to analyse, how human perception decomposes any complex object into primitives and is invariant in nature.

Inspired by the analysis of human cognition, in this work we propose "IPD-Net", which takes in the raw point cloud with Euclidean positions as input and provides an invariant primitive representation. In this work, we represent the point cloud into its primitive shapes, which are planar, cylindrical, conical, and planar. IPD-Net additionally provides an implicit invariant representation, avoiding the need to make the model robust towards rotation with augmentation in the training pipeline. The invariant primitive decomposition of the point cloud with IPD-Net can be used for downstream tasks like classification, segmentation, point cloud completion.

We summarize our contribution as follows:

- We propose IPD-Net: Invariant Primitive Decompositional Network for $SO(3)$ invariant primitive representation of 3D point cloud.
  - We propose to extract Implicit Invariant Features (IIF) towards achieving invariance in decomposition using centric distance field and normals.
  - We propose to extract global signature of the point cloud through Spatial Rectification Unit (SRU) using canonical representation for rotation invariant signature.
- We demonstrate the results of our proposed methodology for $SO(3)$ invariant decomposition on TraceParts Dataset, and compare with state-of-the-art methods.
- We show the generalizability of proposed IPD-Net as plugin for classification of point clouds on benchmark dataset (ModelNet40), and compare with state-of-the-art methods.

In Section 2, we discuss the proposed architecture for extraction of Invariant features towards decomposition of point clouds. We discuss the results and effect of extracted SO(3) invariant features on decomposition and 3D object classification in Section 3 and conclude in Section 4.

2. IPD-Net: SO(3) Invariant Primitive Decomposition Network

We propose IPD-Net: an Invariant Primitive Decomposition Net for $SO(3)$ equivariant point cloud representations. Unlike previous methods, we propose to employ rotation invariant centric distance fields along with per-point normals. We achieve $SO(3)$ equivariant decompositions via Topological Invariant features and canonical representations of proposed centric-distance fields. Towards extracting these feature we propose Implicit Invariant Features (IIF) and Spatial Rectifier Unit (SRU). We fuse the canonical and topological invariant to get higher dimensional representations and project them to 4 primitive shapes probabilities using Shared-MLPs.

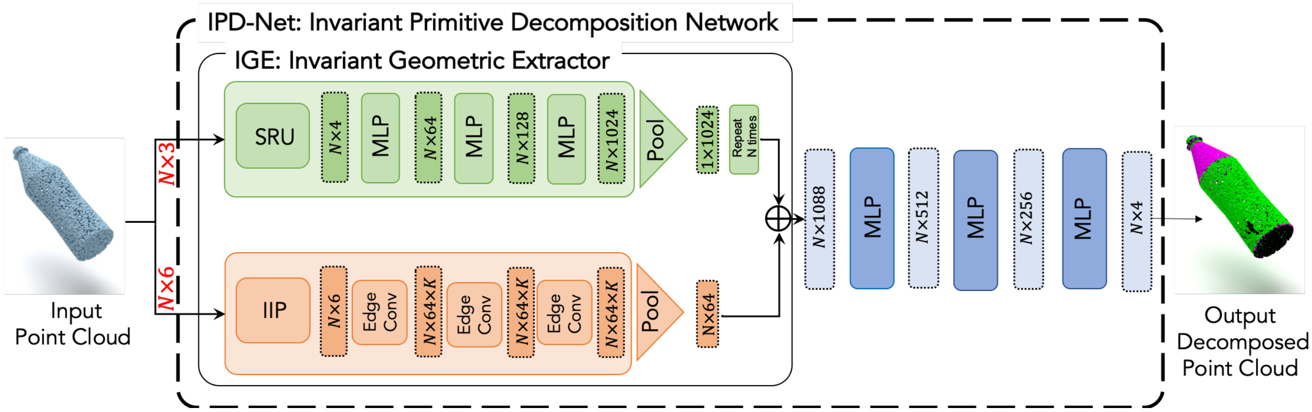

Figure 2. Our architecture takes in point cloud with normals into IGE:Invariant geometric extractor. The IIP: Implicit invariant Projector takes in point and normals to give Invariant Implicit Features (IIF) with which invariant local features using edge conv we do maxpool at K to get per-point feature vector of 64 dimension. The SRU:special Rectifier Unit gives the Canonical representation of to extract global signature consistent with rotations. we concatenate this global relations with every local feature space.Finally with we use shared MLP's to decompose the point cloud into primitives using both local and global features.

2.1. Problem Statement

Let point cloud $P = \{p_1, p_2, ..., p_n\}$ where n represents point density of a given point cloud and $p_i \in \mathbb{R}^6$ containing coordinates (x, y, z) and per-point normals (nx, ny, nz). We propose IPD-Net as a decomposition function f_θ parameterized by weights θ such that it yield a per-point primitive probability distribution $Q = \{q_1, q_2, ..., q_n\}$ belongs to four categories.

We model g_ϕ as per-point invariant feature transformer $g_\phi(RP) = g_\phi(P)$ where R is 3D rotation matrix, making IPD-Net an invariant decomposer and $g_\phi \cup h_\zeta = f_\theta$ as shown in Figure 2.

2.2. Invariant Geometry Extractor (IGE)

Invariant Geometry Extractor block contains Spatial Rectification Unit (SRU) and Implicit Invariant Projector (IIP) modules used to extract SO(3) invariant features implicitly for both global signature and local neighbourhood geometry.

Implicit Invariant Projector (IIP) is the module we use to convert euclidian position into a 6-dimensional vector towards $SO(3)$ invariant features. we modeled an extended daboux frame work to address two crucial challenges in LGR-Net [51]: 1) computation efficiency, and 2) ambiguity with-respect-to the orientation of local topology. We introduce a novel Moment Relativity Field denoted by $\Psi = \mu - P$, where μ represents the moment or centroid of the point cloud coordinates P. $\|\Psi\|$ being a scalar descriptor facilities in understanding the underlying relative typologies of point cloud with-respect-to to centroid. Furthermore, we simplify this descriptor by computing the centroid of the entire point cloud over the centroid of the local group (k-NN) while reducing both computation and memory footprint. By

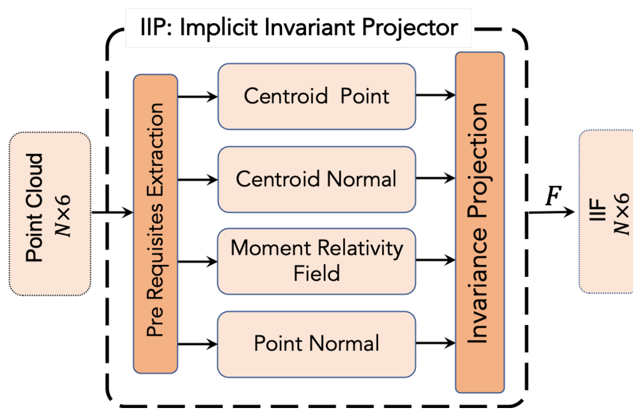

Figure 3. Implicit Invariant Projector takes in the point cloud with normal extracting centroid point, Centroid Normal, Moment Relativity field and point normal to implicitly represent the point cloud using rotation invariant features in a 6-dimensional space called invariant implicit features with which local feature learning takes place.

leveraging the aforementioned descriptor, we derive novel *Implicit Invariant Features* F to incorporate a range of factors given by,

$$F = \Big[\|\Psi\|, \Psi \cdot N, \Psi \cdot N_\mu, N \cdot N_\mu, u(N) \cdot u(N_\mu), v(N) \cdot v(N_\mu) \Big] \tag{1}$$

Here, $u(x) = \Psi \times x$, $v(x) = u(x) \times x$ represent the extended Darboux frame, and $\cdot$ refers to the angle given by dot-product. N represents normals of point cloud P and N_μ represents centroid of normals repeated n times.

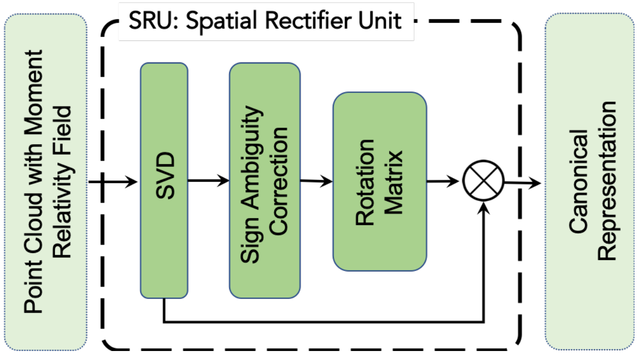

Figure 4. We employ SRU to extract global signatures. We take point cloud with Moment Relativity Field project the point cloud into 4-dimensional canonical space. We achieve robustness in invariance by resolving sign ambiguity. Hence, consistent global signatures irrespective of rotations.

Spatial Rectification Unit (SRU) transforms given set of co-ordinates (x, y, z) to its canonical representation as show in Figure 4. Although implicit invariant features are invariant to rotation, they lack spatial information necessary for downstream tasks like segmentation, upsampling, and reconstruction. Therefore we extract global geometric signature of point cloud using canonical representation. Towards this we employ Singular Value Decomposition (SVD) [12] on point cloud point cloud $P = \{p_1, p_2, ..., p_n\}$ containing (x, y, z) along with our Moment-Relativity Field $\|\Psi\|$ containing geometric descriptions (aiding in better geometric signature) as an additional feature giving us H=[P, $\|\Psi\|$],

$$H = USV^T \tag{2}$$

where, the orthogonal matrix $V^T \in \mathbb{R}^4$. With the the spatial information and distance value for each instance present the orthogonal matrix projects the point cloud into 4-dimensional canonical space where point cloud is aligned uniquely according to its geometry regardless of its initial pose. Thus, helping us in obtaining consistent global signatures.

The rectification through the rotation matrix may contain sign ambiguity, like [9] we resolve the issue with by fixing the sign ambiguity between U and V^T by,

$$l = sgn(U^T, \|\Psi\|) \tag{3}$$

and we apply signs from l to get,

$$V'^T = V^T L \tag{4}$$

where, L is the diagonal matrix with signs l and V'^T is rotation with no sign ambiguity present.

Improved Implicit Features(IIF) we make use of local aggregators (EdgeConv) to extract local geometry. The per point local features learnt are better as IIF inherits geometric descriptors like Moment Relativity Field and normals and since these features are implicitly invariant we conventional feature extractors making it easier to improve with feature learning methodologies.

The canonical representations derived using SRU are used to extract global relation using shared MLP's. We use Moment Relative Field containing better geometric information and get better and unique signatures with geometrical changes. Both the features are concatenated and decomposed into primitive shapes.

3. Results and Discussions

In this section, we discuss about the dataset used for training of our proposed method, experimental setup and setting of proposed methodology, demonstrate the results of our methodology and compare with state-of-the-art methods on decomposition and generalizability on classification with z/z, $z/SO(3)$, and $SO(3)/SO(3)$.

3.1. Datasets

In this section, we discuss on the dataset used for benchmarking of our proposed methodology on decomposition using TraceParts dataset and on classification using ModelNet40.

- **TraceParts [38]**: dataset consists of mechanical component models along with primitive shapes information labeled (planar, spherical, cylindrical and conical) with 12984 training samples and 3173 testing samples.
- **ModelNet40 [47]**: dataset consists of CAD models belonging to 40 categories. These CAD models are sampled to 1024 points to form a pointcloud with 9843 training samples and 2465 testing samples.

3.2. Experimental Setup

In this section, we discuss about the experimental setup of proposed methodology for Invariant Decomposition and Classification.

- **Training Setup for IPD-Net**
  We consider learning rate of 0.001, batch size of 8, using Adam Optimizer and Negative Log likely-hood as a loss function. We consider 1024 points during training using Random sampling. We use augmentation of Random point dropping with 0.15 and random scale with a scaling factor of 0.25. We train IPD-Net just for 50 epochs on TraceParts dataset with all the rotation setups.
- **Training Setup for SO(3) Invariant 3D Classification Setup**
  We consider learning rate of 0.001, batch size of 32, using SGD Optimizer with momentum of 0.9 and

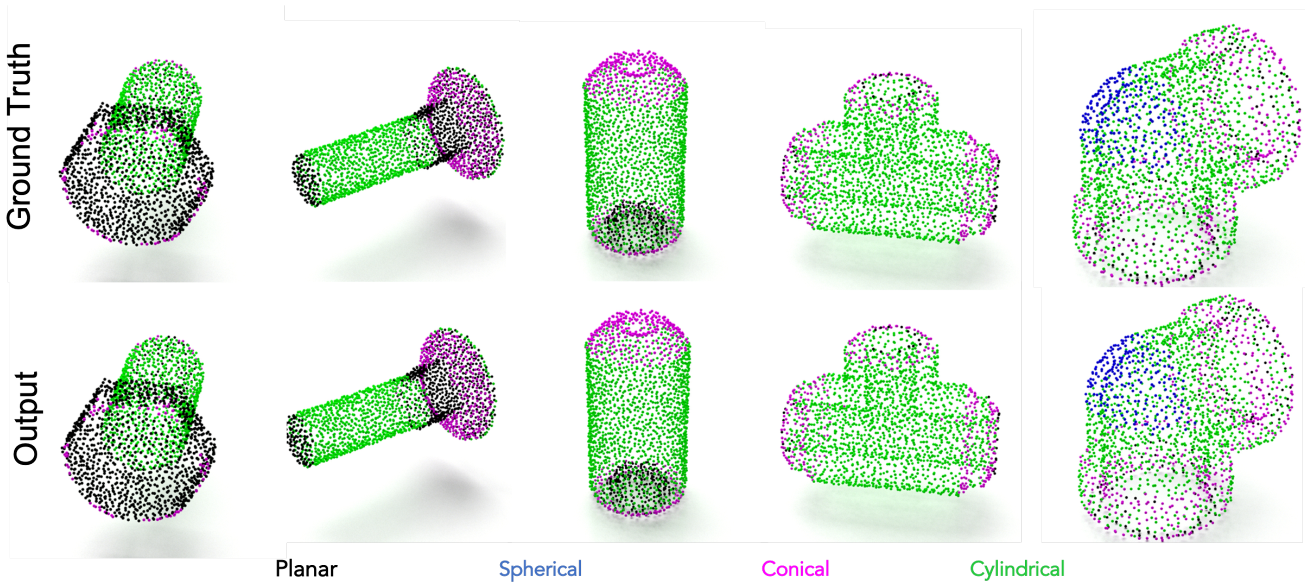

Figure 5. Visualization of decomposition results on TraceParts dataset [38] using our proposed methodology IPD-Net. We infer that the results of IPD-Net are consistent in decomposing all the four primitives shapes and are near to the ground truth. First row represents the Ground truth and second row represents the decomposed outputs of IPD-Net.

Cross Entropy as a loss function. We consider 1024 points during training using Random sampling. We use augmentation of Random point dropping with 0.15 and random scale with a scaling factor of 0.25. We train the classification task for 250 epochs on ModelNet40 dataset with all the rotation setups. We use IPD-Net features to PointNet [33] and Point Cloud Transformer [13].

3.3. Results

In this section, we demonstrate the results of proposed methodology on decomposition and classification and compare it with state-of-the-art methods.

- **Comparison of decomposition results with state-of-the-art methods**:
  We have evaluated our proposed method for rotation-invariant point cloud decomposition on the Traceparts dataset, and the results are presented in Table 1. Our method, IPD-Net, performs less effectively than ABD-Net [17] in both the z/z and $SO(3)/SO(3)$ decomposition tasks. However, we contend that ABD-Net is not invariant in the $z/SO(3)$ setting, as evidenced by its significant drop in performance of 59.99% from z/z to $z/SO(3)$ in mean Intersection over Union of Decompositions. On the other hand, IPD-Net exhibits stable and robust performance, maintaining 0% decrease in performance in the same setting.

Table 1. The decomposition accuracy of proposed methodology on TraceParts dataset [38] in comparison with state-of-the-art method ABD-Net [17]. We compare the results of IPD-net on different settings of rotation z/z, $z/SO(3)$, and $SO(3)/SO(3)$. Here **Bold** represents the best performance; ↑ represents higher is better and ↓ represents lower is better. **mIoU** refers to Mean of Intersection over Union and **sIoU** refers to Standard deviation of Intersection over Union.

	z/z		$z/SO(3)$		$SO(3)/SO(3)$	
	mIoU ↑	**sIoU** ↓	**mIoU** ↑	**sIoU** ↓	**mIoU** ↑	**sIoU** ↓
ABD-Net [17](2021)	**0.9552**	**0.0355**	0.3607	0.1537	**0.9216**	**0.0734**
IPD-Net (Ours)	0.9028	0.1045	**0.9028**	**0.1045**	0.9028	0.1045

- **Qualitative Analysis of decomposition results with state-of-the-art methods**:
  To highlight the efficacy of our proposed method IPD-Net, we show qualitative comparison of our decompositions against ABD-Net [17]. Figure 7 shows visual supremacy of IPD-Net over ABD-Net in all point clouds across all $SO(3)$ rotation as justified in Table 1. Both model were trained for $z/SO(3)$ setting, we visually depict and highlight some of the few limitation of ABD-Net. When we randomly rotation excluding z axis the input point cloud. ABD-Net fails in Identifying screen of Laptop as 'Planar', body of airplane and rocket as 'Cylindrical' as shown in highlighted regions in Figure 7. Where as IPD-Net is robust in Identifying screen of Laptop as Planar, body of airplane and rocket as cylindrical. One potential reason to this may be due to incorporation of our proposed implicit invariant features. We also report that ABD-Net is robust to

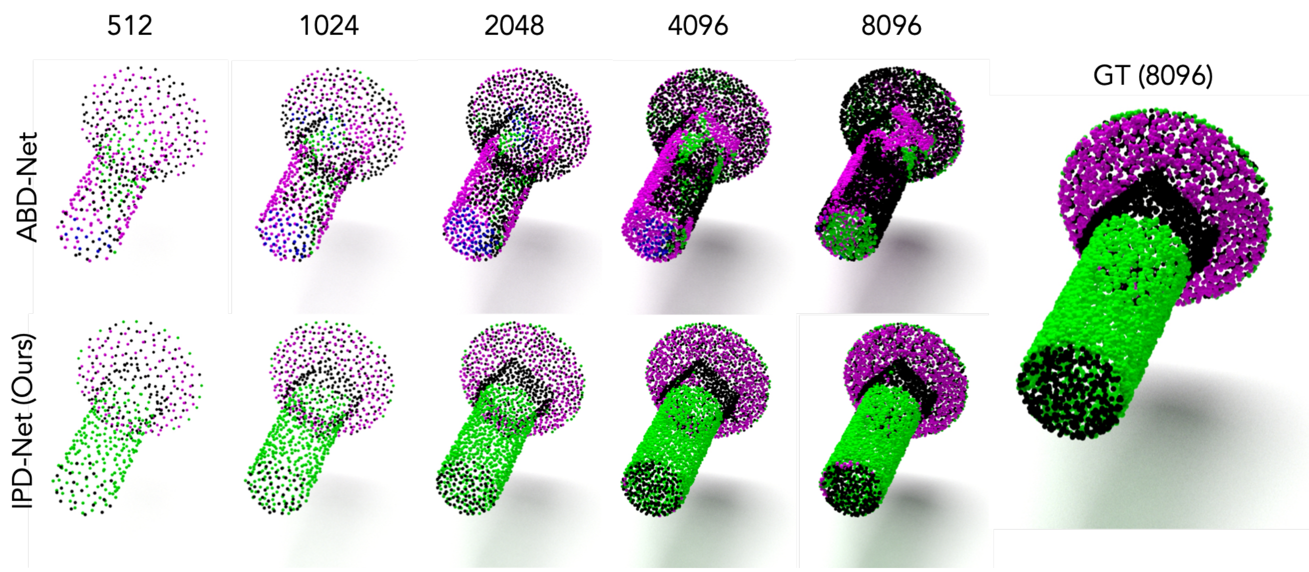

Figure 6. Visualization of decomposition of varying points of point cloud using IPD-Net and ABD-Net [17]. Number of points varying from 512, 1024, 2048, 8096. IPD-Net and ABD-Net [17] are trained on 1024 points. IPD-Net outperforms ABD-Net [17] in-terms to robustness for density varying point clouds.

the rotations that is available during training (i.e, z-axis) as depicted in highlighted region of Stool, where it accurately identifies the cylindrical geometry which is oriented with-respect-to z-axis.

Unlike ABD-Net [17] which is susceptible to new farthest point sampling pattern of point cloud that was not available for training. The susceptiblity of ABDNet towards surfacial pertubation is explored in Figure 6.

- **Comparison of SO(3) invariant classification with state-of-the-art methods**: We have assessed the performance of our proposed method, IPD-Net, as a plug-in for rotation-invariant point cloud classification and compared it with other state-of-the-art methods, as presented in Table 2. Our method serves as a plug-and-play module for existing non-invariant methods, as explained in Section 3.2. One of our IPD variants, PCT [13], achieved the third-highest accuracy on the ModelNet40 dataset [47] for the $SO(3)$-invariant point cloud classification task, outperforming other rotation-invariant methods.

4. Conclusions

In this paper, we have proposed IPD-Net: Invariant Primitive Decompositional Network for $SO(3)$ invariant primitive representation of 3D point cloud. Towards this, we have proposed to extract Implicit Invariant Features (IIF) to achieve invariance in decomposition using Moment Relative Field and normals. Towards extracting canonical representation for rotation invariant global signature of the point cloud, we have proposed Spatial Rectification Unit

Table 2. The classification accuracy of proposed methodology in comparison with state-of-the-art methods on ModelNet40 with 1024 point density. We plug our invariant features proposed methodology with PointNet [33] and PCT [13]. We demonstrate the classification results on different settings of rotation z/z, $z/SO(3)$, and $SO(3)/SO(3)$. Highest values are represented in **Bold**, second highest values are represented in Underline and the third highest values are represented in **Bold and Underlined** format.

	z/z	$z/SO(3)$	$SO(3)/SO(3)$
PointNet [33](2016)	85.9	19.6	74.7
PointNet++ [34](2017)	91.8	28.4	85.0
Spherical-CNN [7](2017)	88.9	76.7	86.9
PCNN [1](2018)	92.3	11.9	85.1
PointCNN [21](2018)	92.5	**41.2**	84.5
ShellNet [50](2019)	93.1	19.9	87.8
a^3S-CNN [22](2019)	89.6	**87.9**	88.7
PCT [13](2021)	87.3	24.6	87.3
Robust Methods			
TFN [43](2018)	88.5	85.3	87.6
SFCNN [35](2019)	91.4	84.8	90.1
RI-Conv [48](2019)	86.5	86.4	86.4
SPHNet [32](2019)	87.7	86.6	87.6
ClusterNet [5](2019)	87.1	87.1	87.1
GC-Conv [49](2020)	89.0	89.1	89.2
RI-GCN [19](2020)	91.0	**91.0**	91.0
LGR-Net [51](2022)	90.9	90.9	91.1
IPD+PointNet(Ours)	87.7	87.7	87.7
IPD+PCT(Ours)	89.3	**89.3**	89.3

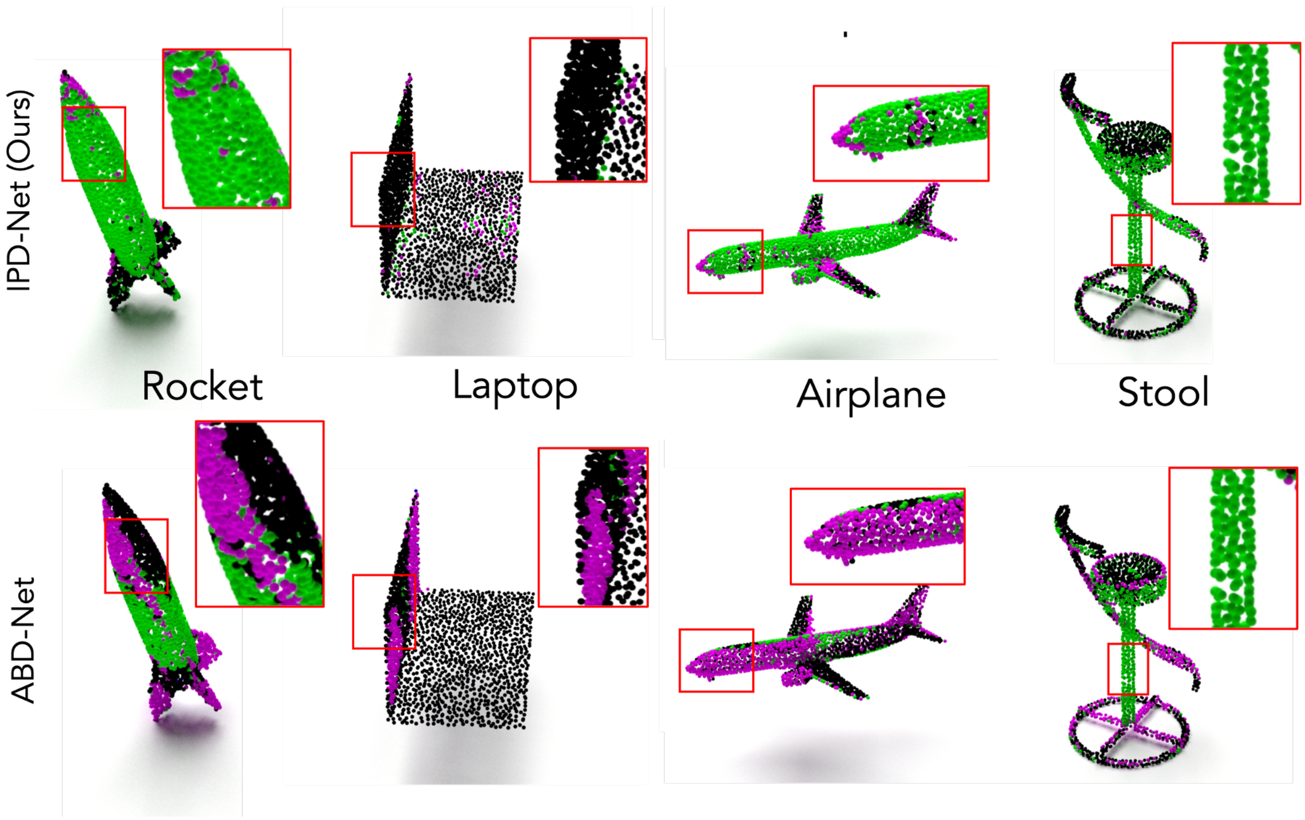

Figure 7. Visual comparison between our proposed point cloud primitive decomposition method, IPD-Net, and the state-of-the-art method ABD-Net [17] on four objects: Rocket, Laptop, Airplane, and Stool. Both models were trained for the $z/SO(3)$ setting, and we highlight some limitations of ABD-Net [17] when the input point cloud is randomly rotated excluding the z axis. Specifically, ABD-Net [17] fails to identify the screen of the Laptop as *Planar* and the body of the Airplane and Rocket as *Cylindrical*, as shown in the highlighted regions. In contrast, IPD-Net is robust in identifying these primitive shapes. We speculate that the incorporation of our proposed implicit invariant features contributes to this robustness. Additionally, we report that ABD-Net [17] performs well when the objects are oriented with respect to the z axis only, as evidenced by the accurate identification of the cylindrical geometry of the Stool in the highlighted region.

(SRU). We have demonstrated the results of our proposed methodology for $SO(3)$ invariant decomposition on TraceParts Dataset, and have compared the results of decomposition with state-of-the-art methods. While our proposed method may not outperform ABD-Net in certain decomposition tasks, we have shown that ABD-Net is not invariant in the $z/SO(3)$ setting, resulting in a significant drop in performance of 59.99% from z/z to $z/SO(3)$ in mean Intersection over Union of Decompositions. In contrast, IPD-Net maintains stable and robust performance in this setting, with a 0% decrease in performance. We have shown the generalizability of proposed IPD-Net as plugin for classification of point clouds on benchmark dataset (ModelNet40), and have compared with state-of-the-art methods.

5. Acknowledgement

This work is partly carried out under the AICTE-RPS project "Shape Representation, Reconstruction and Rendering of 3D Models" (8-247/ RIFD/ RPS (POLICY-r)/2OLB-19) and Department of Science and Technology (DST) through the ICPS programme - Indian Heritage in Digital Space for the project "Crowd-Sourcing" (DST/ ICPS/ IHDS/ 2018 (General)).

References

[1] Matan Atzmon, Haggai Maron, and Yaron Lipman. Point convolutional neural networks by extension operators. *ACM Trans. Graph.*, 37(4), jul 2018. 6

[2] Irving Biederman. Human image understanding: Recent research and a theory. *Computer Vision, Graphics, and Image Processing*, 32(1):29–73, 1985. 2

[3] Irving Biederman. Recognition-by-components: a theory of human image understanding. *Psychological review*, 94 2:115–147, 1987. 2

[4] Thomas O. Binford. Visual perception by computer. In *Proceedings of the IEEE Conference on Systems and Control (Miami, FL)*, 1971. 2

[5] Chao Chen, Guanbin Li, Ruijia Xu, Tianshui Chen, Meng Wang, and Liang Lin. Clusternet: Deep hierarchical cluster network with rigorously rotation-invariant representation for point cloud analysis. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4989–4997, 2019. 6

[6] Kevin Ellis, Daniel Ritchie, Armando Solar-Lezama, and Joshua B. Tenenbaum. Learning to infer graphics programs from hand-drawn images. *CoRR*, abs/1707.09627, 2017. 2

[7] Carlos Esteves, Christine Allen-Blanchette, Ameesh Makadia, and Kostas Daniilidis. Learning so(3) equivariant representations with spherical cnns. *CoRR*, 2017. 6

[8] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, jun 1981. 2

[9] Kent Fujiwara, Ikuro Sato, Mitsuru Ambai, Yuichi Yoshida, and Yoshiaki Sakakura. Canonical and compact point cloud representation for shape classification. *CoRR*, abs/1809.04820, 2018. 4

[10] Syed Altaf Ganihar, Shreyas Joshi, Shankar Setty, and Uma Mudenagudi. 3d object decomposition and super resolution. In *SIGGRAPH Asia 2014 Posters*, SA '14, New York, NY, USA, 2014. Association for Computing Machinery. 2

[11] Syed Altaf Ganihar, Shreyas Joshi, Shankar Shetty, and Uma Mudenagudi. Metric tensor and christoffel symbols based 3d object categorization. In *ACM SIGGRAPH 2014 Posters*, SIGGRAPH '14, New York, NY, USA, 2014. Association for Computing Machinery. 2

[12] G. H. Golub and C. Reinsch. Singular value decomposition and least squares solutions. *Numerische Mathematik*, 14(5):403–420, Apr 1970. 4

[13] Meng-Hao Guo, Jun-Xiong Cai, Zheng-Ning Liu, Tai-Jiang Mu, Ralph R. Martin, and Shi-Min Hu. Pct: Point cloud transformer. *Computational Visual Media*, 7(2):187–199, Apr 2021. 5, 6

[14] Deepti Hegde, Dikshit Hegde, Ramesh Ashok Tabib, and Uma Mudenagudi. Relocalization of camera in a 3d map on memory restricted devices. In R. Venkatesh Babu, Mahadeva Prasanna, and Vinay P. Namboodiri, editors, *Computer Vision, Pattern Recognition, Image Processing, and Graphics*, pages 548–557, Singapore, 2020. Springer Singapore. 1

[15] Eldar Insafutdinov and Alexey Dosovitskiy. Unsupervised learning of shape and pose with differentiable point clouds. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018. 2

[16] Sudhakaran Jain and Hamidreza Kasaei. 3d_den: Open-ended 3d object recognition using dynamically expandable networks. *IEEE Transactions on Cognitive and Developmental Systems*, 2021. 2

[17] Siddharth Katageri, Shashidhar V Kudari, Akshaykumar Gunari, Ramesh Ashok Tabib, and Uma Mudenagudi. Abd-net: Attention based decomposition network for 3d point cloud decomposition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pages 2049–2057, October 2021. 2, 5, 6, 7

[18] Siddharth Katageri, Sameer Kulmi, Ramesh Ashok Tabib, and Uma Mudenagudi. Pointdccnet: 3d object categorization network using point cloud decomposition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 2200–2208, June 2021. 2

[19] Seohyun Kim, Jaeyoo Park, and Bohyung Han. Rotation-invariant local-to-global representation learning for 3d point cloud. *CoRR*, abs/2010.03318, 2020. 6

[20] Lingxiao Li, Minhyuk Sung, Anastasia Dubrovina, Li Yi, and Leonidas J. Guibas. Supervised fitting of geometric primitives to 3d point clouds. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2647–2655, 2019. 2

[21] Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. Pointcnn: Convolution on x-transformed points. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. 6

[22] Min Liu, Fupin Yao, Chiho Choi, Sinha Ayan, and Karthik Ramani. Deep learning 3d shapes using alt-az anisotropic 2-sphere convolution. In *International Conference on Learning Representations*, 2019. 6

[23] Nenad Markuš, Igor Pandžić, and Jörgen Ahlberg. Learning local descriptors by optimizing the keypoint-correspondence criterion: Applications to face matching, learning from unlabeled videos and 3d-shape retrieval. *IEEE Transactions on Image Processing*, 28(1):279–290, 2019. 2

[24] Jiri Matas and O Chum. Randomized ransac with td,d test. *Image and Vision Computing*, 22:837–842, 09 2004. 2

[25] Pierre Moulon, Pascal Monasse, Romuald Perrot, and Renaud Marlet. OpenMVG: Open multiple view geometry. In *International Workshop on Reproducible Research in Pattern Recognition*, pages 60–74. Springer, 2016. 2

[26] Raúl Mur-Artal and Juan D. Tardós. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Transactions on Robotics*, 33(5):1255–1262, 2017. 1

[27] Sudhagar Nagarajan, Ishwarya Srikanth, Satarupa Khamaru, and Madasamy Arockiasamy. Imaging and laser scanning–based noncontact deflection monitoring technique for timber railroad bridges. *Practice Periodical on Structural Design and Construction*, 28(1):04022072, 2023. 2

[28] Shanthika Naik, Uma Mudenagudi, Ramesh Tabib, and Adarsh Jamadandi. Featurenet: Upsampling of point cloud and it's associated features. In *SIGGRAPH Asia 2020 Posters*, SA '20, New York, NY, USA, 2020. Association for Computing Machinery. 1

[29] Chengjie Niu, Jun Li, and Kai Xu. Im2struct: Recovering 3d shape structure from a single rgb image. 04 2018. 2

[30] Despoina Paschalidou, Ali Osman Ulusoy, and Andreas Geiger. Superquadrics revisited: Learning 3d shape parsing beyond cuboids. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2019. 2

[31] Alex Pentland. Parts: Structured descriptions of shape. In *AAAI Conference on Artificial Intelligence*, 1986. 2

[32] A. Poulenard, M. Rakotosaona, Y. Ponty, and M. Ovsjanikov. Effective rotation-invariant point cnn with spherical harmonics kernels. In *2019 International Conference on 3D Vision (3DV)*, pages 47–56, Los Alamitos, CA, USA, sep 2019. IEEE Computer Society. 6

[33] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*, 2016. 2, 5, 6

[34] Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*, 2017. 2, 6

[35] Yongming Rao, Jiwen Lu, and Jie Zhou. Spherical fractal convolutional neural networks for point cloud recognition. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 452–460, 2019. 6

[36] Lawrence G. Roberts. Machine perception of three-dimensional solids. In *Outstanding Dissertations in the Computer Sciences*, 1963. 2

[37] T. Santoshkumar, Deepti Hegde, Ramesh Ashok Tabib, and Uma Mudenagudi. Refining sfm reconstructed models of indian heritage sites. In *SIGGRAPH Asia 2020 Posters*, SA '20, New York, NY, USA, 2020. Association for Computing Machinery. 1

[38] TraceParts S.A.S. 4, 5

[39] Shankar Setty, Syed Altaf Ganihar, and Uma Mudenagudi. Framework for 3d object hole filling. In *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, pages 1–4, 2015. 1

[40] Gopal Sharma, Difan Liu, Subhransu Maji, Evangelos Kalogerakis, Siddhartha Chaudhuri, and Radomír Měch. Parsenet: A parametric surface fitting network for 3d point clouds. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision – ECCV 2020*, pages 261–276, Cham, 2020. Springer International Publishing. 2

[41] Hang Su, Subhransu Maji, Evangelos Kalogerakis, and Erik Learned-Miller. Multi-view convolutional neural networks for 3d shape recognition. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015. 1

[42] Ramesh Ashok Tabib, Yashaswini V. Jadhav, Swathi Tegginkeri, Kiran Gani, Chaitra Desai, Ujwala Patil, and Uma Mudenagudi. Learning-based hole detection in 3d point cloud towards hole filling. *Procedia Computer Science*, 171:475–482, 2020. Third International Conference on Computing and Network Communications (CoCoNet'19). 1

[43] Nathaniel Thomas, Tess E. Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds. *CoRR*, abs/1802.08219, 2018. 6

[44] P.H.S. Torr and A. Zisserman. Mlesac: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156, 2000. 2

[45] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics (TOG)*, 2019. 2

[46] Huanshu Wei, Zhijian Qiao, Zhe Liu, Chuanzhe Suo, Peng Yin, Yueling Shen, Haoang Li, and Hesheng Wang. End-to-end 3d point cloud learning for registration task using virtual correspondences. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2678–2683, 2020. 2

[47] Zhirong Wu, Shuran Song, Aditya Khosla, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets for 2.5d object recognition and next-best-view prediction. *CoRR*, abs/1406.5670, 2014. 4, 6

[48] Zhiyuan Zhang, Binh-Son Hua, David W. Rosen, and Sai-Kit Yeung. Rotation invariant convolutions for 3d point clouds deep learning. *CoRR*, abs/1908.06297, 2019. 6

[49] Zhiyuan Zhang, Binh-Son Hua, Wei Chen, Yibin Tian, and Sai-Kit Yeung. Global context aware convolutions for 3d point cloud understanding. pages 210–219, 11 2020. 6

[50] Zhiyuan Zhang, Binh-Son Hua, and Sai-Kit Yeung. Shellnet: Efficient point cloud convolutional neural networks using concentric shells statistics. pages 1607–1616, 10 2019. 6

[51] Chen Zhao, Jiaqi Yang, Xin Xiong, Angfan Zhu, Zhiguo Cao, and Xin Li. Rotation invariant point cloud analysis: Where local geometry meets global topology. *Pattern Recognition*, 127:108626, 2022. 3, 6

[52] Yin Zhou and Oncel Tuzel. Voxelnet: End-to-end learning for point cloud based 3d object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 1