This CVPR workshop paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

VG-VAE: A Venatus Geometry Point-Cloud Variational Auto-Encoder

Tejas Anvekar, Ramesh Ashok Tabib, Dikshit Hegde, Uma Mudengudi
Center of Excellence in Visual Intelligence (CEVI), KLE Technological University.
Vidya Nagar, Hubballi, Karnakata, India.
anvekartejas@gmail.com, ramesh_t@kletech.ac.in, dikshithegde@gmail.com, uma@kletech.ac.in

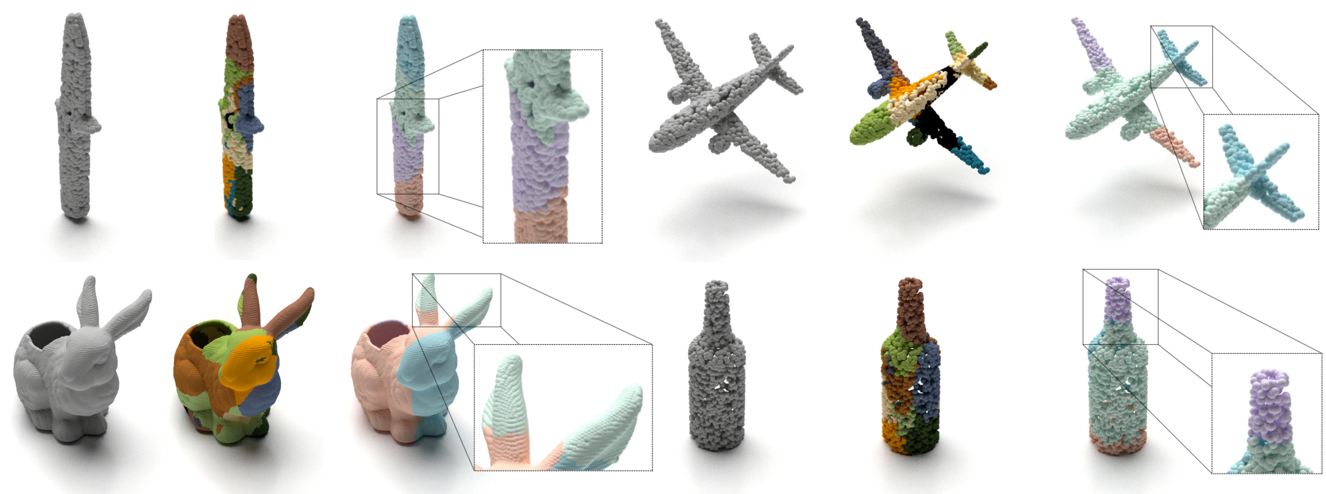

Figure 1. **VG-VAE** is a variational immanent method to learn Hierarchical intrinsic, local and global geometric signatures. Geometric signatures aid in the comprehension of pointcloud morphology. Understanding morphology is equivalent to unearthing Poincloud's creation, and is observed in this Figure. **Grey** represents input pointcloud, the next **Coloured** pointcloud represents geometric signatures $\mathbb{G}$, and the highlighted pointcloud represents geometric labels $\mathcal{G}$.

Abstract

In this paper, we propose VG-VAE: Venatus Geometric Variational Auto-Encoder for capturing unsupervised hierarchical local and global geometric signatures in pointcloud. Recent research emphasises the significance of the underlying intrinsic geometry for pointcloud processing. Our contribution is to extract and analyse the morphology of the pointcloud using the proposed Geometric Proximity Correlator (GPC) and variational sampling of the latent. The extraction of local geometric signatures is facilitated by the GPC, whereas the extraction of global geometry is facilitated by variational sampling. Furthermore, we apply a naive mix of vector algebra and 3D geometry to extract the basic per-point geometric signature, which assists the unsupervised hypothesis. We provide statistical analyses of local and global geometric signatures. The impacts of our geometric features are demonstrated on pointcloud classification as downstream task using the classic pointcloud feature extractor PointNet. We demonstrate our analysis on ModelNet40 a benchmark dataset, and compare with state-of-the-art techniques.

1. Introduction

In recent years, 3D pointcloud has grown in relevance in a variety of fields, including 3D printing, Metaverse, and self-driving automobiles as sensors capture depth in addition with other visual signals. Pointclouds are beginning to play a crucial role in many real-world applications, such as simultaneous localization and mapping (SLAM) [14] [5], 3D object detection, digitization of heritage sites towards presentation in AR/VR/XR/MR. Towards presentation of 3D data in digital space, there is a need for efficient methods to analyze, process and derive huge volume of 3D pointclouds [15] [20] [25] [22] [23]. In contrast to human vision, supervising a machine to infer geometric information

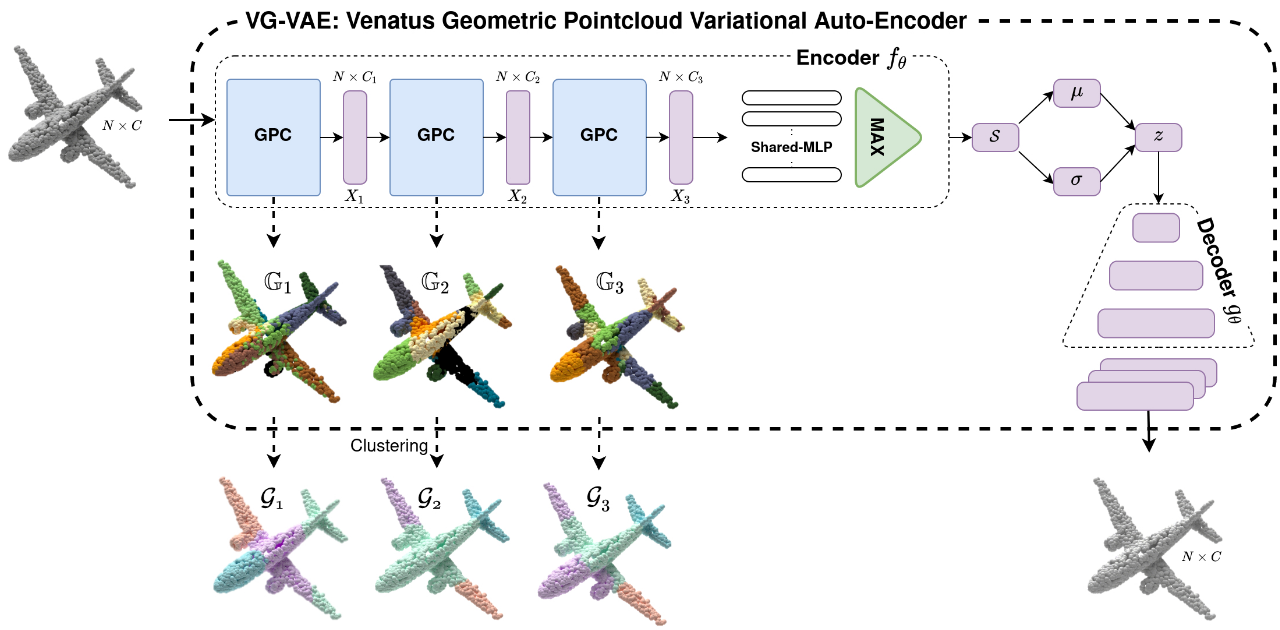

Figure 2. The proposed framework VG-VAE: Venatus Geometry Point-Cloud Variational Auto-Encoder. Here GPC represents Geometric Proximity Correlator, $\boldsymbol{X_i}$ represents Geometric features, $\mathbb{G}_i$ represents Geometric signatures and $\boldsymbol{\mathcal{G}_i}$ represents Geometric labels, $\boldsymbol{\mathcal{S}}$ represents empirical latent space, $\boldsymbol{\mu}$ represents empirical mean, $\boldsymbol{\sigma}$ represents empirical log variance and $\boldsymbol{z}$ represents sampled latent space.

is challenging. Representing 3D object as a collection of intrinsic structure provides morphology as a geometric signature. This abbreviated representation of a 3D object is critical for improving shape understanding, processing, and analysis. To address the lack of geometric information in 3D pointclouds, various studies attempt to decompose a 3D pointcloud into meaningful structures, and infer a topological graph by modelling correlations between these parts [16] [9] [10] as geometric signatures. However, the decomposition is completely perceptual. While, some approaches attempt to recover a parametric form of the input pointclouds using defined geometric shapes [12] [24] [9] [10].

Authors in [10] [9] discuss about deriving fundamental geometric signatures and use them as a plugin towards downstream tasks. These techniques find challenges generalizing across datasets and hybrid geometries. Towards this, we propose self-supervised approach to eliminate the need for generalization across datasets, and generate redesigned hybrid geometric signatures for a chosen dataset unlike [10] [9].

The main contributions of the work are:

- We propose a self supervised / unsupervised technique to derive geometric features of pointcloud and name it as VG-VAE: Venatus Geometry Point-Cloud Variational Auto-Encoder.
  - A novel Geometric Proximity Correlator (GPC) to co-relate intrinsic and empirical geometric signatures of a pointcloud.
    - ∗ An Intrinsic Geometric Interpreter (IGI) to derive the features of pointcloud in existing and higher dimensional space.
    - ∗ A KNN prior based geometric feature aware encoder.
- We propose to consider variational sampling to generalize over extrinsic geometric features across the samples.
- We model a geometry aware pointcloud classifier using the derived geometric features as a plugin towards classification of pointclouds.
- We demonstrate the effect of derived geometric signatures on benchmark dataset (ModelNet40) and compare the results with state-of-the-art techniques.

In Section 2 we discuss the proposed VG-VAE: Venatus Geometry Point-Cloud Variational Auto-Encoder. We discuss the ablation and results of proposed model and compare with state-of-the-art techniques in Section 3, and conclude in Section 4.

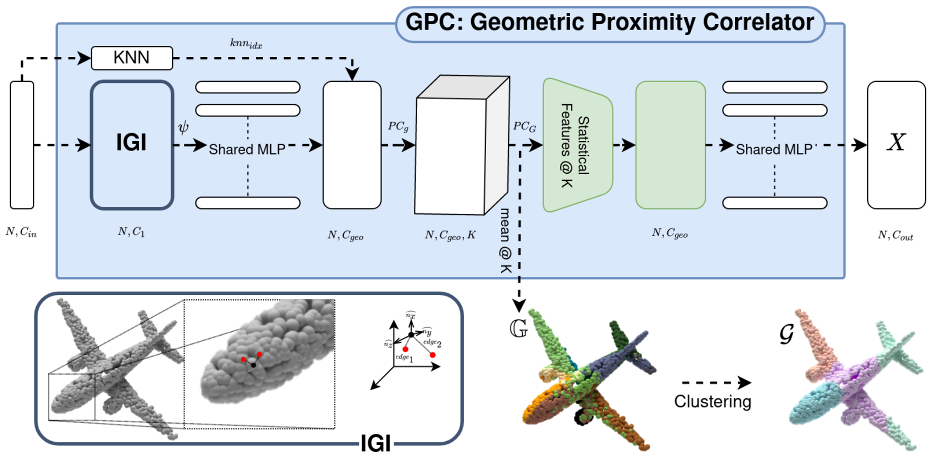

Figure 3. Isoloted view of GPC: Geometric Proximity Correlator that extracts the individual geometric signatures. Here IGI: Intrinsic Geometry Interpreter facilitates extraction of intrinsics.

2. Venatus Geometry Point-Cloud Variational Auto-Encoder (VG-VAE)

We model **VG-VAE**, Venatus Geometry Point-Cloud Variational Auto-Encoder for learning unsupervised hierarchical geometric features in pointcloud data. Geometric aspects of a pointcloud characterise its morphology which facilitates abstraction, generation, and description of the pointcloud as shown in Figure 2. Towards this, we propose a Geometric Proximity Correlator (GPC) for extracting hierarchical local and global geometric signatures. Further we estimate the underlying per-point variations in higher dimensional space with nearest neighbour as a prior using Variational sampling.

2.1. Geometric Proximity Correlator (GPC)

Geometric Proximity Correlator facilitates to capture the geometric signatures. Geometric signatures are computed via a seven step process as shown in Algorithm 1. Initially, we search for the prior i.e., indices $\boldsymbol{idx_{knn}}$ using K-Nearest neighbours. Intrinsic Geometry Interpreter ψ is estimated on the given pointcloud $\boldsymbol{pc}$. Towards extracting the empirical (learnt) features $\boldsymbol{PC_G}$, we emphasis on weight shared mlp that projects ψ to a higher dimensional space. The geometric posterior $\boldsymbol{X}$ is captured by gathering the nearest neighbour $\boldsymbol{gathered_{neighbours}}$ in higher dimensional space aided by prior $\boldsymbol{idx_{knn}}$ as shown in Figure 3. To boost the learning, we compute the statistical features (Mean, Standard deviation, Min, and Max) of $\boldsymbol{gathered_{neighbours}}$ with respect to its K nearest neighbours and project them to higher dimensional space using a weight shared mlp. The proposed geometric signatures $\mathbb{G}$ are captured by tapping $\boldsymbol{PC_{\mathbb{G}}}$ and computing mean for the nearest neighbours, whereas the geometric labels $\boldsymbol{\mathcal{G}}$ are simply obtained by computing clusters $\mathbb{G}$. The idiosyncratic clusters are inspired from Attention Based Decomposition network [9] and PointDCCNet [10].

Algorithm 1: Geometric Proximity Correlator

Input: Point Cloud $\rightarrow$ $\boldsymbol{PC}$; // $\boldsymbol{B, N, C}$

Output: Geometric Proximity Correlator's weights $\boldsymbol{f_\theta}$; Geometry aware pointcloud features $\boldsymbol{X}$; Geometric Signatures $\mathbb{G}$ and Geometric Cluster labels $\boldsymbol{\mathcal{G}}$.

Initialize $\boldsymbol{K}$, k, $\boldsymbol{f_\theta}$, $\boldsymbol{C_{in}}$, $\boldsymbol{C_{geo}}$, and $\boldsymbol{C_{out}}$ according to Section 2.

$\boldsymbol{idx_{knn}}$ = Knn index($\boldsymbol{PC}$,$\boldsymbol{K}$).

/* $\boldsymbol{B, N, C}$ */

ψ = intrinsic geometry interpreter ($\boldsymbol{PC}$).

/* $\boldsymbol{B, N, C_1}$ */

$\boldsymbol{PC_G}$ = shared mlp($\boldsymbol{C_{in}} + \boldsymbol{C_1}$, $\boldsymbol{C_{geo}}$)(ψ).

/* $\boldsymbol{B, N, C_{geo}}$ */

$\boldsymbol{gathered_{neighbors}}$ = gather operation($\boldsymbol{idx_{knn}}$, $\boldsymbol{PC_G}$).

/* $\boldsymbol{B, N, C_{geo}, K}$ */

min, max, std, and mean $\leftarrow$ $\boldsymbol{gathered_{neighbors}}$.

$\boldsymbol{X}$ $\leftarrow$ shared mlp($\boldsymbol{C_{geo}}$, $\boldsymbol{C_{out}}$)(Concat(min, max,std,and mean)).

/* $\boldsymbol{B, N, C_{out}}$ */

$\mathbb{G}$ = mean($\boldsymbol{PC_G}$, dim=-1).

$\boldsymbol{\mathcal{G}}$ = KMeans(k, $\mathbb{G}$).

2.1.1 Intrinsic Geometry Interpreter (IGI)

The unsupervised hypothesis of our network is notably aided by joint learning, variational sampling of latent and a basic Intrinsic Geometry Interpreter ψ. The variational sampling of latent guarantees that the global signatures of the pointcloud is captured, whilst the geometric interpreter enables encapsulating the local signatures. The Intrinsic Geometry Interpreter ψ is a naive combination of vector algebra and 3D geometric as shown in Figure 3. Pointclouds are widely used to produce meshes by triangulation. By forming a triangle with respect to two nearest neighbors (pc_{j1}, pc_{j2}) of pointcloud pc_i, we estimate Intrinsic Geometry Interpreter ψ_i by:

$$\psi_i = \begin{cases} pc_i = x, y, z; & pc_i \in \mathbb{R}^n, \\ \overrightarrow{v_1} = pc_{j1} - pc_i; & \overrightarrow{v_1} \in \mathbb{R}^n, \\ \overrightarrow{v_2} = pc_{j2} - pc_i; & \overrightarrow{v_2} \in \mathbb{R}^n, \\ |\overrightarrow{v_1}| = \ell_2(\overrightarrow{v_1}); & |\overrightarrow{v_1}| \in \mathbb{R}^1, \\ |\overrightarrow{v_2}| = \ell_2(\overrightarrow{v_2}); & |\overrightarrow{v_2}| \in \mathbb{R}^1, \\ \hat{n} = \overrightarrow{v_1} \times \overrightarrow{v_2}; & \hat{n} \in \mathbb{R}^3. \end{cases} \tag{1}$$

where $(\overrightarrow{v_1}, \overrightarrow{v_2})$ represents edge (1, 2) respectively of given point pc_i, $(|\overrightarrow{v_1}|, |\overrightarrow{v_2}|)$ represents edge lengths and $\hat{n}$ represent normals of pointcloud pc_i. Unlike Geometric Backpropagation Net [19], we propose utilizing ψ_i beyond three dimensions, with the exception of excluding $\hat{n}$, as the cross-product has the orthogonality characteristic only in three and seven dimensional spaces. The set of per-point geometry aids in the robustness of our hypothesis, since normals and edge length would immediately reveal outliers in the supplied point set.

Towards interpretability and hierarchical learning of local to global geometric signatures, we propose to utilize Geometric Proximity Correlator in a dolly chain fashion multiple times. The intermediate $\mathbb{G}_i$'s and $\mathcal{G}_i$'s correspond to local to global geometric signatures.

2.2. Variational Sampling

The hierarchical global symmetricity $\mathcal{S}$ is captured by pooling the global geometric features X through a weight shared mlp inspired by PointNet [17]. The variational sample space z captures the extrinsic geometry of the pointcloud space through reparametrization technique and ELBOW loss [11].

Variational Autoencoders (VAE) are generative models capable of learning approximated data distributions via variational inference. We consider the stochastic latent space $\mathbf{z}$ and optimize the upper-bound on the negative log-likelihood of $\mathbf{x}$:

$$\mathbb{E}_{\mathbf{x}\sim p_d(\mathbf{x})}[-\log p(\mathbf{x})] < \mathbb{E}_{\mathbf{x}}[\mathbb{E}_{\mathbf{z}\sim q(\mathbf{z}|\mathbf{x})}] + \mathbb{E}_{\mathbf{x}}\Big[\mathcal{D}\Big]$$

$$\mathcal{D} = KL\big(q(\mathbf{z}|\mathbf{x})||p(\mathbf{z})\big)$$

where $KL(\cdot||\cdot)$ is Kullback-Leibler divergence [8], $p_d(\mathbf{x})$ is empirical distribution, $q(\mathbf{z}|\mathbf{x})$ is variational posterior (the encoder), $p(\mathbf{x}|\mathbf{z})$ is generative model (the generator) and $p(\mathbf{z})$ is the prior. Sampling of $q(\mathbf{z}|\mathbf{x})$ is performed via reparametrization trick facilitated by optimizing a neural network's weights for obtaining $p(\mathbf{x}|\mathbf{z})$ and $q(\mathbf{z}|\mathbf{x})$. Total loss used to tune VAE are represented by two terms reconstructions loss (ℓ_2 norm taken from difference between sampled and reconstructed) such that $p(\mathbf{x}|\mathbf{z})$ is assumed to be normal-distribution and regularization term that compels $\mathbf{z}$ to be produced from $q(\mathbf{z}|\mathbf{x})$ from a prior distribution $p(\mathbf{z})$.

We emphasise on variational sampling for two main reasons

- To generalize the extrinsic geometric signatures over the data distribution unlike Vanilla Autoencoder.
- Towards generations of synthetic pointcloud by sampling the geometric aware feature space unlike 3D-AAE [31].

Towards tuning of geometric features by optimizing the weights of encoder f_θ and decoder g_θ as shown in Figure 2, we propose to utilize chamfer distance as reconstruction loss given by,

Chamfer pseudo Distance (**CD**): measures the squared two norm ℓ_2^2 between each point in one set S_1 to its nearest neighbour in the other set S_2 as shown below.

$$CD\,(S_1, S_2) = \sum_{x\in S_1} \min_{y\in S_2} \|x-y\|_2^2 + \sum_{y\in S_2} \min_{x\in S_1} \|x-y\|_2^2 \tag{2}$$

and KL Divergence as a regularizor.

$$KL(p,q) = \log\frac{\sigma}{\sigma_e} + \frac{\sigma^2 + (\mu_e - \mu)^2}{2\sigma^2} - \frac{1}{2} \tag{3}$$

where (μ, σ) are empirical mean and standard deviation as shown in Figure 2. (μ_e, σ_e) are expected mean and standard deviation for a given pointcloud.

The proposed decoder is feature extrapolation net comprised of mlp's similar to an image based VAE [11].

2.3. Pointcloud Classification

We incorporate the produced unsupervised geometric signatures $\mathbb{G}$ to the input pointcloud and feed it through the PointNet architecture to demonstrate the impact of our proposed methodology. We demonstrate our analysis of the impact made by hierarchical local and global geometric signatures on ModelNet40 [30] a benchmark dataset and compare them with state-of-the-art techniques.

3. Results and Discussions

In this section, we discuss about the dataset used, evaluation metrics and comparison of our methodology with state-of-the-art methods.

3.1. Datasets

To evaluate the performance of the proposed methodology we use benchmark ModelNet40 dataset and compare the classification accuracy with state-of-the-art methods PointNet [17], PointNet++ [18], KCNet [6], MRTNet [3], SpecGCN [27], DGCN [28], PCNN [1], ABDNet [9], PointConv [29]. PointDCCNet [10], PointTransformer [32], RSCNN [13], PCT [4]. We compare our geometrical results with ABDNet [9].

- **ModelNet 40 [30]**: dataset consists of CAD models belonging to 40 categories. These CAD models are sampled to 2048 points to form a pointcloud.
- **ShapeNet [2]**: ShapeNet is a single clean 3D models and manually verified category and alignment annotations. It covers 55 common object categories with about 51,300 unique 3D models.

3.2. Experimentation Details

In this section, we discuss the experimental setup as two major parts.

- **Architectural Details**: The details for architecture of VG-VAE, GPC, and variational sampling decoder are shown in Table 1. To make it more understandable table is colour coded with respect Figure 2. The architecture of GPC is elaborated in Algorithm 1.
- **Training Setup VG-VAE**: Initially during training the weights of VG-VAE's Encoder f_{θ} and Decoder g_{θ} set to uniform distribution and expected mean μ_e and standard deviation σ_e from Equation 3, is set to (0, 0.02) respectively. The weights θ of VG-VAE are tuned using proposed variational loss given in Equation 3 and chamfer pseudo distance as reconstruction loss given in Equation 1. The variational sampling is explained in Section 2.2. We train VG-VAE for 2000 epochs until KL divergence has reached to minima. The intermediate geometric signatures $\mathbb{G}_i$ are saved for computing geometric labels $\mathcal{G}_i$ using KMeans clustering. We find number of cluster K = 4 yields better results visually and also facilitates pointcloud classification.
- **Training Setup VG-VAE + PointNet**: We train a geometry aware 3D classifier by merging geometric features of VG-VAE with pointcloud, we concatenate the pointcloud and geometric feature and train PointNet, the hyperparameters for classification are set as per PointNet [17] with an exception of replacing in channels of $\mathbf{1}^{st}$ layer to dimensional of geometric features and pointcloud together. During training VG-VAE's weights are freezed so that the discriminative loss wont affect extracted geometric features.

Table 1. Architectural Details of VG-VAE. Here represents Geometric Proximity Correlator (GPC) to co-relate intrinsic and empirical (learnt) geometric signature and represents variational sampling over extrinsic geometric features.

Type	Configs	#	Keys
GPC	$(C_{in}, C_{geo}, C_{out}, K)$	(14,12,64,32)	GPC 1
GPC	$(C_{in}, C_{geo}, C_{out}, K)$	(194,12,128,64)	GPC 2
GPC	$(C_{in}, C_{geo}, C_{out}, K)$	(386,12,512,128)	GPC 3
Conv1d	(C_{in}, C_{out})	(512,512)	Shared MLP
Linear	(in, out)	(512,1024)	μ, σ
Sampling	NA	NA	z
Linear	(in, out)	(1024,128)	-
Linear	(in, out)	(128,512)	-
Linear	(in, out)	(512,1024)	-
Linear	(in, out)	(1024,2048*3)	pc

Table 2. The classification accuracy of proposed methodology with comparison with state-of-the-art method on ModelNet40 with 1024 point density. 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

Methods	Input	Accuracy
PointNet [17](2017)	xyz	89.2
PointNet++ [18](2017)	xyz	90.7
MRTNet [3](2018)	xyz	91.2
Spec-GCN [27](2018)	xyz	91.5
Spec-GCN [27](2018)	xyz, nor	91.8
PCNN [1](2018)	xyz	92.3
PointConv [29](2018)	xyz,nor	92.5
DGCNN [28](2019)	xyz	92.2
RSCNN [13](2019)	xyz	**92.9**
3D-GCNN [26](2019)	xyz	83.5
Point Transformer [32](2020)	xyz	**93.7**
Point Transformer [32](2020)	xyz,nor	92.8
KCNet [6](2021)	xyz	91.0
PointDCCnet [10](2021)	xyz,nor	92.5
ABD-Net [9](2021)	xyz	92.2
ABD-Net [9](2021)	xyz,nor	92.8
PCT [4](2021)	xyz	93.2
Ours VG-VAE + PointNet	xyz	**92.9**
Ours VG-VAE + Linear SVM	xyz	84.5

3.3. Results

In this section, we analyse the inference of our experiments. Initially we discuss the implications of unsupervised hierarchical geometrical signatures as shown in Figure 1. We can infer that the Geometrical Signatures $\mathbb{G}$ of VG-VAE change due to unsupervised hypothesis. Towards this, we propose to the cluster the Geometrical signatures $\mathbb{G}$ to attain inferable Geometrical Labels $\mathcal{G}$ as shown in highlighted regions in Figure 1. The Geometrical labels are the obtained from KMeans [7], where cluster indexes are computed using distance rather than density. The Geometric

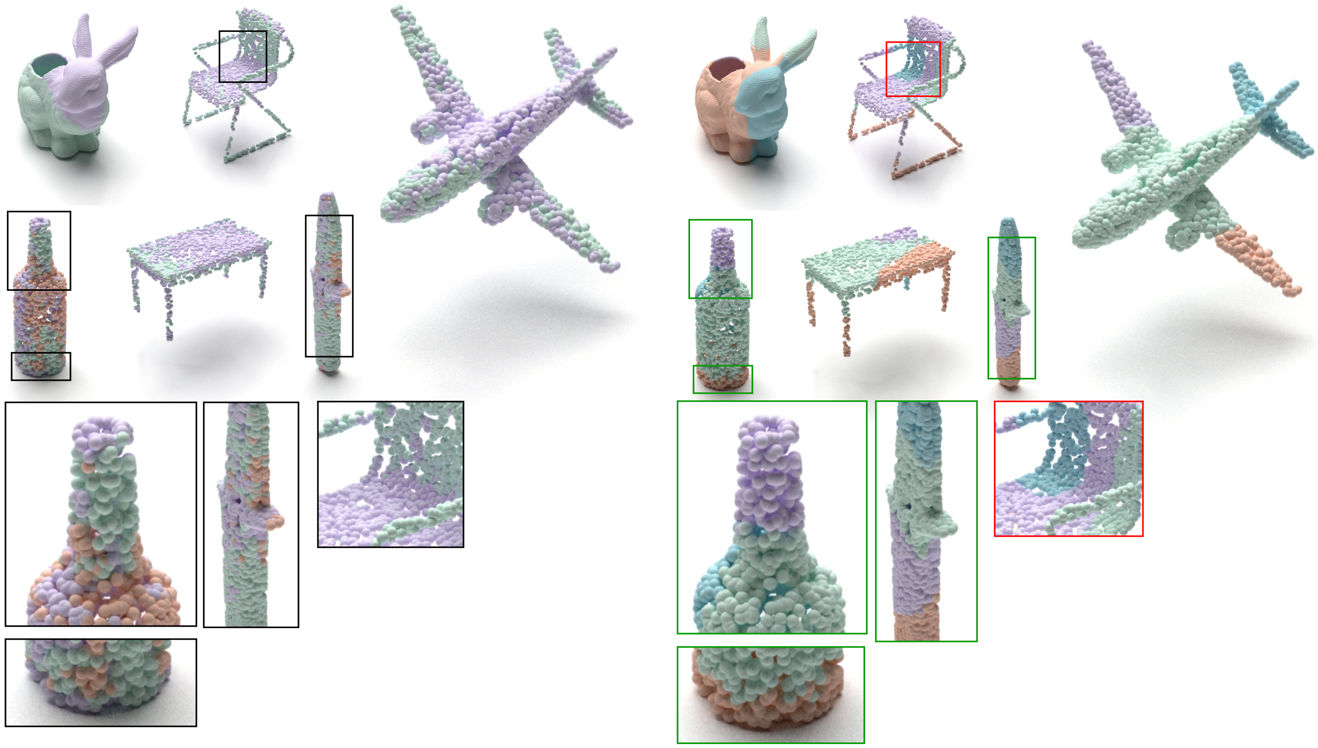

Figure 4. Visual supremacy of our proposed **VG-VAE**'s unsupervised geometric labels $\mathcal{G}_3$ achieved over supervised Attention Based Decomposition Network (ABD-Net) labels. Left side shows the results of ABD-Net and VG-VAE towards right side. We highlighted area of the given pointclouds (*"Bottle, Knife and Chair"*) at the specific area with maximal change in geometric signatures. We compare ours against ABD-Net therefore green colour boxes represents better performance of VG-VAE over ABD-Net and red colour indicates vice-versa.

Features of VG-VAE utilise the property of KMeans, as $\mathbb{G}$ is distance function in a higher dimensional space.

We compare our geometric features with state-of-the-art decomposition network ABD-Net [9]. We find that supervised methodology do not generalize to capture geometric signatures across various datasets. The issue lies with overfitting of supervised methods on a single dataset, and finds challenges with other data distribution. We show similar scenario where, ABD-Net achieve state-of-the-art 99% decomposition accuracy on ANSI [21] dataset, but fail to capture basic geometric signature, when tested on ShapeNet [2]. In Figure 4, we analyse the performance of both algorithms. We compare ours against ABD-Net therefore green colour highlights in Figure 4, represents better performance of VG-VAE over ABD-Net and red colour indicates vice-versa. By considering results of ABD-Net as reference, as it is supervised method, we highlight specific regions of 6 pointclouds (*"Bunny", "Chair", "Bottle", "Table", "Knife", and "Airplane"*) where maximal geometric change could be observed. Comparing ABD-Net we observe:

- VG-VAE captures better Geometric Signatures (*"Tail","Wings", and "Engine""*) on Airplane as shown in Figure 4, compared to ABD-Net.
- On observing Bottle VG-VAE is able to capture the change in geometry at bottle-head and at base of bottle, unlike ABD-Net.
- Even comparing Knife ABD-Net fails to capture the required geometric features unlike VG-VAE.
- Comparing Table, Both the methods yield similar results.
- On Chair ABD-Net is more generalized compared VG-VAE in capturing flat surfaces.
- On Bunny its hard to infer, tho VG-VAE captures change in geometries at ears.

Comprehensively VG-VAE is more robust in capturing geometrical features. We do not perform evaluation on obtained features as they are unsupervised trained.

We evaluate the effect of geometric signature by merging these as a clue to PointNet classifier and a Linear SVM. We compare the classification accuracy of our method with state-of-the-art supervised pointcloud classification method as shown in Table 2. We can infer that our method compared to discriminative methods such as ABD-Net and PointDC-CNet, our method outperforms the both.

The experiments are conducted on NVIDIA RTX 3090 GPU with 24GB RAM and AMD RYZEN threadripper 3970x CPU using PyTorch framework.

4. Conclusions

In this work, we have proposed VG-VAE: Venatus Geometric Variational Auto-Encoder for capturing unsupervised hierarchical local and global geometric signatures in pointcloud. The extraction of local geometric signatures is facilitated by the Geometric Proximity Correlator, whereas the extraction of global geometry is facilitated by variational sampling. We have provided statistical analyses of local and global geometric signatures. The impacts of our geometric features are demonstrated on pointcloud classification as downstream task using the classic pointcloud feature extractor PointNet. We have demonstrated our analysis on ModelNet40 a benchmark dataset and compare them with state-of-the-art techniques.

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. *CoRR*, abs/1803.10091, 2018. 5

[2] Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. ShapeNet: An Information-Rich 3D Model Repository. Technical Report arXiv:1512.03012 [cs.GR], Stanford University — Princeton University — Toyota Technological Institute at Chicago, 2015. 5, 6

[3] Matheus Gadelha, Rui Wang, and Subhransu Maji. Multiresolution tree networks for 3d point cloud processing. In *ECCV*, 2018. 5

[4] 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

[5] 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

[6] Jinyung Hong and Theodore P. Pavlic. Kcnet: An insect-inspired single-hidden-layer neural network with randomized binary weights for prediction and classification tasks. *CoRR*, abs/2108.07554, 2021. 5

[7] Xin Jin and Jiawei Han. *K-Means Clustering*, pages 563–564. Springer US, Boston, MA, 2010. 5

[8] James M. Joyce. *Kullback-Leibler Divergence*, pages 720–722. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. 4

[9] 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, 3, 5, 6

[10] 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, 3, 5

[11] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. 2013. 4

[12] 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

[13] Yongcheng Liu, Bin Fan, Shiming Xiang, and Chunhong Pan. Relation-shape convolutional neural network for point cloud analysis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8895–8904, 2019. 5

[14] Raul Mur-Artal and Juan D. Tardós. ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras. *CoRR*, abs/1610.06475, 2016. 1

[15] 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

[16] Xiaojuan Ning, Er Li, Xiaopeng Zhang, and Yinghui Wang. Shape decomposition and understanding of point cloud objects based on perceptual information. VRCAI '10, page 199–206, New York, NY, USA, 2010. Association for Computing Machinery. 2

[17] Charles R. Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017. 4, 5

[18] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on

point sets in a metric space. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. 5

[19] Shi Qiu, Saeed Anwar, and Nick Barnes. Geometric back-projection network for point cloud classification. *IEEE Transactions on Multimedia*, 2021. 4

[20] 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

[21] TraceParts S.A.S. 6

[22] 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

[23] Shankar Setty, Himanshu Shekhar, and Uma Mudenagudi. Region of interest (roi) based 3d inpainting. In *SIGGRAPH ASIA 2016 Posters*, SA '16, New York, NY, USA, 2016. Association for Computing Machinery. 1

[24] 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

[25] 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. 1

[26] Rishabh Tigadoli, Ramesh Ashok Tabib, Adarsh Jamadandi, and Uma Mudenagudi. 3d-gcnn - 3d object classification using 3d grid convolutional neural networks. pages 269–276, 2019. 5

[27] Chu Wang, Babak Samari, and Kaleem Siddiqi. Local spectral graph convolution for point set feature learning. *arXiv preprint arXiv:1803.05827*, 2018. 5

[28] 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. 5

[29] Wenxuan Wu, Zhongang Qi, and Li Fuxin. Pointconv: Deep convolutional networks on 3d point clouds. *arXiv preprint arXiv:1811.07246*, 2018. 5

[30] 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, 5

[31] Maciej Zamorski, Maciej Zieba, Rafal Nowak, Wojciech Stokowiec, and Tomasz Trzcinski. Adversarial autoencoders for generating 3d point clouds. *CoRR*, abs/1811.07605, 2018. 4

[32] Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip H. S. Torr, and Vladlen Koltun. Point transformer. *CoRR*, abs/2012.09164, 2020. 5