TreeGCN-ED: A Tree-Structured Graph-Based Autoencoder Framework for Point Cloud Processing

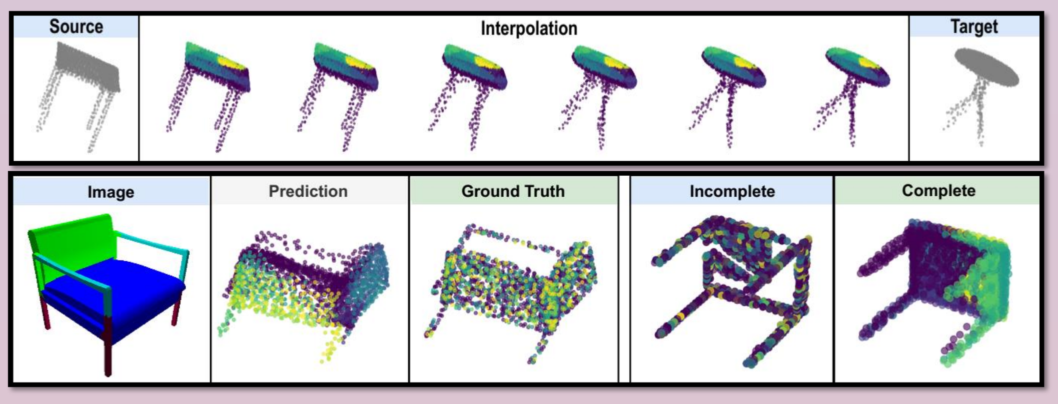
Prajwal Singh, Ashish Tiwari, Kaustubh Sadekar, and Shanmuganathan Raman

Computer Vision, Imaging, and Graphics Lab
Indian Institute of Technology Gandhinagar

PG23 Daejeon, Korea

PROBLEM

- We present a tree-structured graph-based autoencoder framework to generate robust embeddings of point clouds through hierarchical information aggregation.
- The learned embeddings are discriminative enough to distinguish among different object classes.
- They are robust enough for applications such as point cloud clustering, interpolation, completion, and single image-based point cloud reconstruction.



RELATED WORK

Several methods have been proposed for encoding-decoding of point cloud data such as PointNet, VoxelNet, PointCNN, and PointRCNN. PointNet [1] and FoldingNet [2] frameworks encode the point cloud to lower-dimensional embeddings carrying rich information about the point clouds and being used for downstream tasks like clustering and classification over point clouds. Recently in TreeGAN [3], the authors have proposed a tree-structured decoder which uses the idea of graph convolution to generate a point cloud using a noise vector $\boldsymbol{z} \in \mathbb{R}^{96}$ sampled from a normal distribution $N(0, I)$. It aggregates the information from parent nodes at each layer instead of spatially adjacent nodes to leverage the tree-structured decoder architecture when applying graph convolution. Our work develops around restructuring TreeGAN to an autoencoder for learnig robust point cloud embeddings.

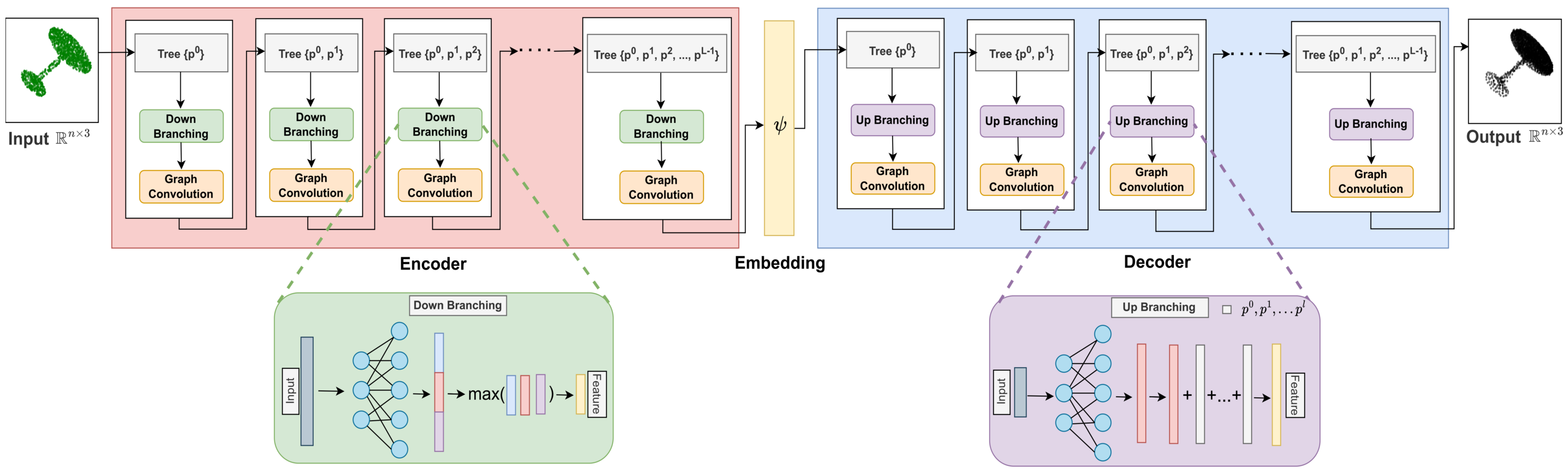
OVERVIEW

- Due to the irregular structure of point clouds compared to images and 3D voxels, it is challenging to design autoencoders to learn rich embeddings over point clouds.
- We design a deep encoder-decoder framework to learn information-rich robust embeddings for several tasks on point clouds, such as clustering, classification, interpolation, completion, and image-based reconstruction.
- We extend Tree-GAN [3] - a GAN-based tree-structured framework to TreeGCN-ED, an autoencoder based framework, to efficiently encode and decode point clouds.
- Specifically, we develop an encoder to generate robust embedding of point clouds that are then used by TreeGAN decoder to generate point clouds.

CVIG
IIT Gandhinagar

METHODOLOGY

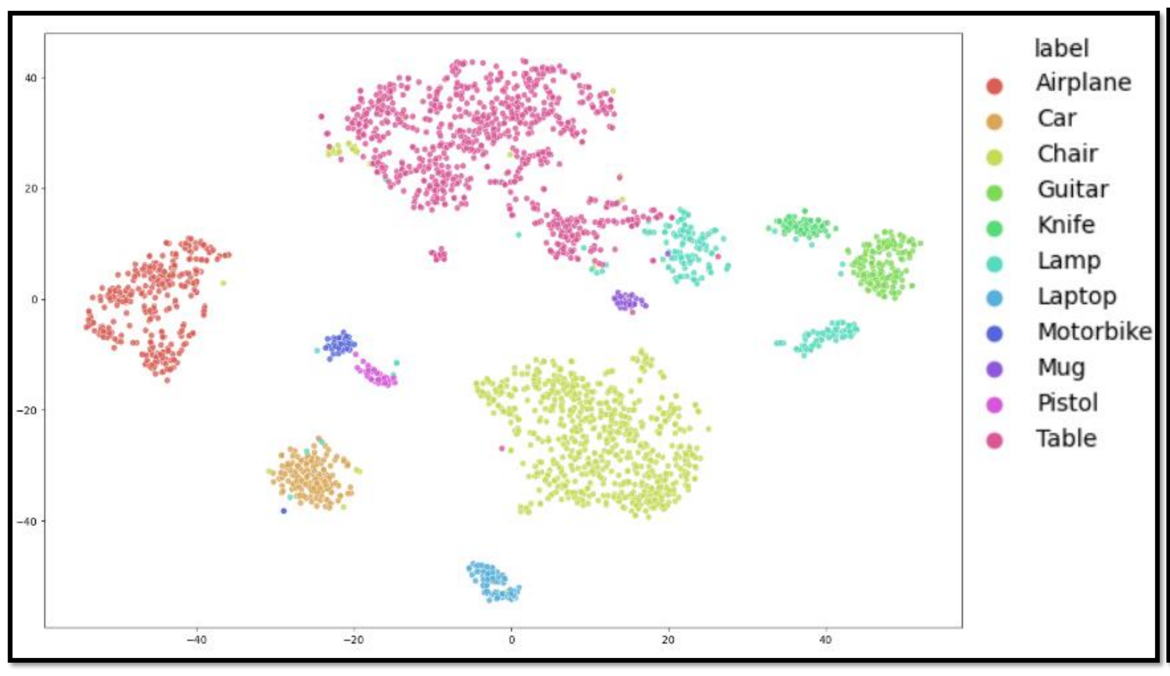
- The encoder (left) consists of multiple stages of down-branching and graph convolutions for encoding the input 3D point cloud $p \in \mathbb{R}^{n\times3}$ into a feature embedding $\psi \in \mathbb{R}^{K}$.
- The decoder (right) takes ψ as input and reconstructs the 3D point cloud through a set of up-branching and graph convolutions (similar to [3]).
- The down-branching consists of a fully-connected layer followed by max-pooling to accumulate features from ancestors for each node. The output of the fully-connected layer is divided into C equal components that are passed to the max-pooling layer.



- The up-branching is responsible for collecting information from the feature embedding of the ancestors and upsampling, which are then passed to the graph convolution layer for further refinement.
- We use Chamfer loss to train the network over the ShapeNetBenchmarkV0 [4] dataset with 16 object classes. We perform train-validation-test split as per [3].

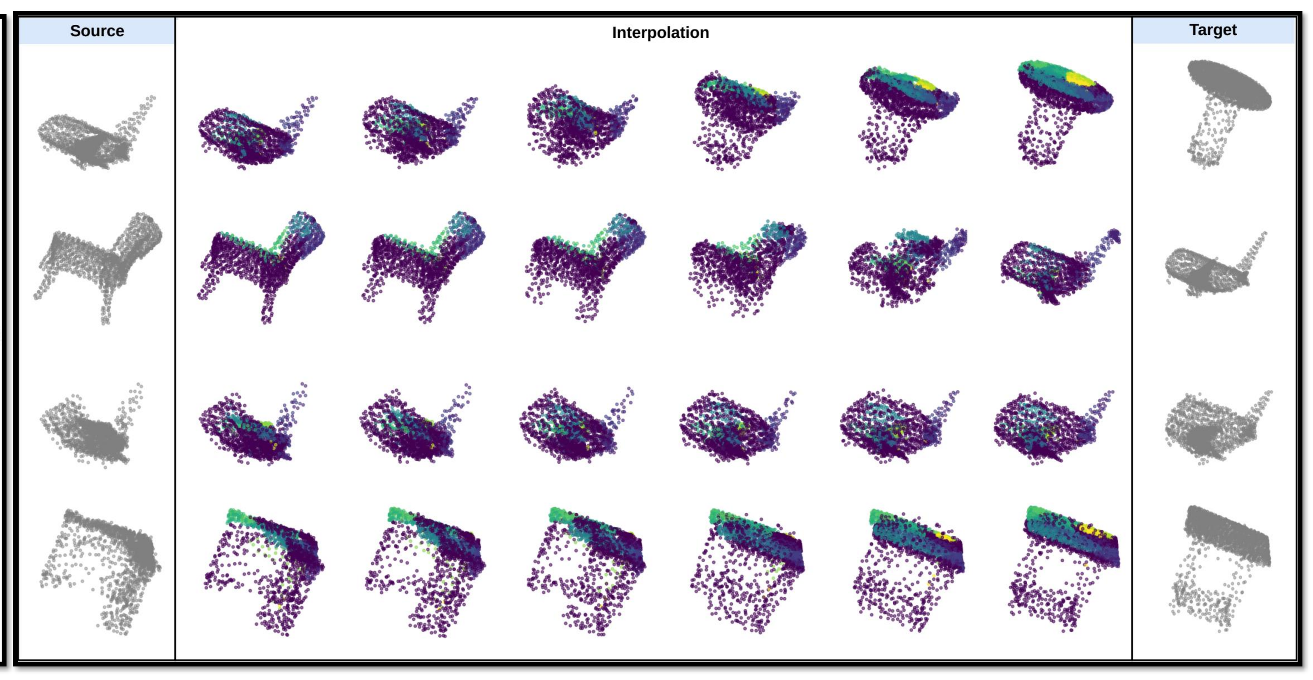
RESULTS

(a) Point Cloud Clustering



t-SNE plot to establish how well our encoder model can generate feature embedding for each class. The inter-class separation is higher, indicating high discriminative capacity. Perplexity value = 40)

(b) Point Cloud Interpolation



Inter-class (top) and intra-class (bottom) point cloud interpolation exhibiting a smooth transition. Results illustrate the ability of our model to synthesize novel shapes between two given shapes and faithfully represent the object class at each interpolation stage.

(c) Point Cloud Completion and Single Image-Based Point Cloud Reconstruction

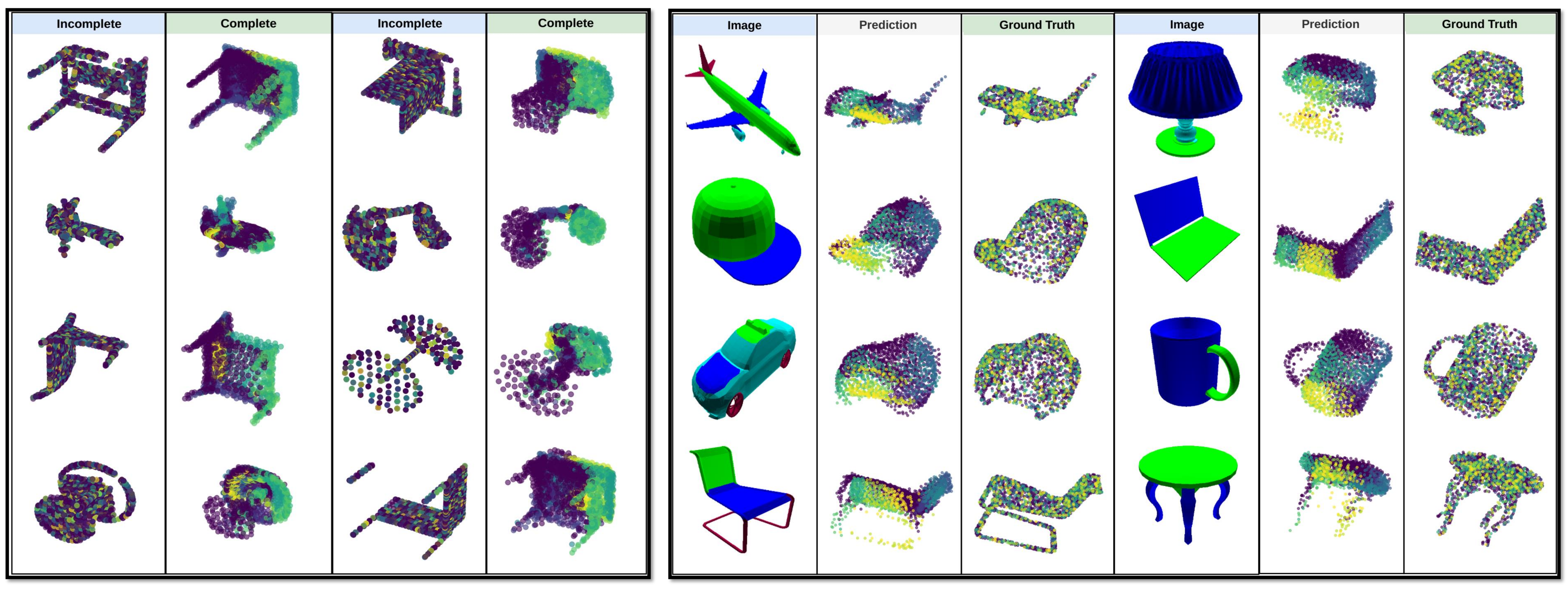


Figure shows how well TreeGCN-ED can fill the missing structures in the point clouds (left) and that the learned embeddings can also foster image-based reconstruction (right)

REFERENCES

[1] Qi, Charles R., et al. "Pointnet: Deep learning on point sets for 3d classification and segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[2] Yang, Yaoqing, et al. "Foldingnet: Point cloud auto-encoder via deep grid deformation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
[3] Shu, Dong Wook, Sung Woo Park, and Junseok Kwon. "3d point cloud generative adversarial network based on tree structured graph convolutions." Proceedings of the IEEE/CVF international conference on computer vision. 2019.
[4] Chang, Angel X., et al. "Shapenet: An information-rich 3d model repository." arXiv preprint arXiv:1512.03012 (2015).