

PU-GCN: Point Cloud Upsampling using Graph Convolutional Networks

Guocheng Qian*

Abdullah Abualshour*

Guohao Li

Ali Thabet

Bernard Ghanem

King Abdullah University of Science and Technology (KAUST)

{guocheng.qian, abdullah.abualshour, guohao.li, ali.thabet, bernard.ghanem} @kaust.edu.sa

Abstract

The effectiveness of learning-based point cloud upsampling pipelines heavily relies on the upsampling modules and feature extractors used therein. For the point upsampling module, we propose a novel model called **NodeShuffle**, which uses a Graph Convolutional Network (GCN) to better encode local point information from point neighborhoods. **NodeShuffle** is versatile and can be incorporated into any point cloud upsampling pipeline. Extensive experiments show how **NodeShuffle** consistently improves state-of-the-art upsampling methods. For feature extraction, we also propose a new multi-scale point feature extractor, called **Inception DenseGCN**. By aggregating features at multiple scales, this feature extractor enables further performance gain in the final upsampled point clouds. We combine **Inception DenseGCN** with **NodeShuffle** into a new point upsampling pipeline called **PU-GCN**. **PU-GCN** sets new state-of-art performance with much fewer parameters and more efficient inference. Our code is publicly available at <https://github.com/guochengqian/PU-GCN>.

1. Introduction

Point clouds are a popular way to represent 3D data. This increasing popularity stems from the increased availability of 3D sensors like LiDAR. Such sensors are now a critical part of important applications in robotics and self-driving cars. However, due to hardware and computational constraints, these 3D sensors often produce sparse and noisy point clouds, which show evident limitations especially for small objects or those far away from the camera. Therefore, point cloud upsampling, the task of converting sparse, incomplete, and noisy point clouds into dense, complete, and clean ones, is attracting much attention.

Following the success in image super-resolution [4, 17, 12, 23], deep learning methods now achieve state-of-the-art results in point cloud upsampling [38, 39, 37, 16]. Most

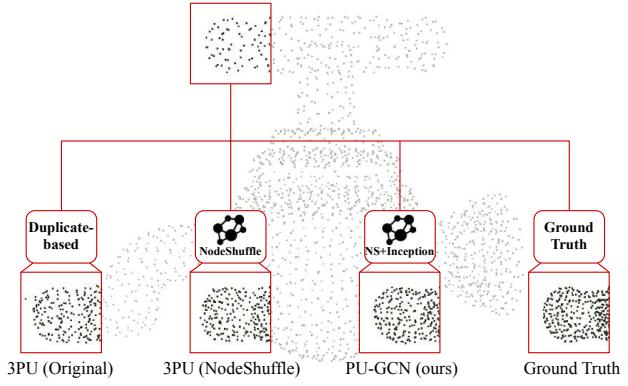


Figure 1: Effectiveness of proposed **NodeShuffle and **Inception DenseGCN**.** We propose a Graph Convolutional Network (GCN) based upsampling module **NodeShuffle** and a multi-scale feature extractor **Inception DenseGCN**. Integrating **NodeShuffle** into the 3PU [37] upsampling pipeline allows for better upsampling and better structure preservation capability. We propose PU-GCN that combines both **Inception DenseGCN** and **NodeShuffle** (NS) upsampling modules. In PU-GCN, **Inception DenseGCN** can further improve upsampling quality and generate fine-grained details (*e.g.* the neck and ball shape of the faucet). The original 3PU uses duplicate-based upsampling.

deep upsampling pipelines comprise two major components: feature extraction and point upsampling. The performance of the point upsampling component tends to define the effectiveness of the final network. Current methods use either multi-branch MLPs (PU-Net [39]) or a duplicate-based approach (3PU [37] and PU-GAN [16]) to upsample 3D points. Multi-branch MLPs operate on each point separately, ignoring any neighborhood information, while duplicate upsampling methods tend to generate point patches similar to the input point clouds. Although the feature extraction modules used in their networks can encode the locality, these shortcomings of the upsampling modules still lead to upsampled point clouds that lack local detail (see

* Guocheng and Abdallah contributed equally to this work.

Figure 1). To better represent locality and aggregate the point neighborhood information, we leverage the power of graphs and specifically Graph Convolutional Networks (GCNs). GCNs are considered a powerful tool to process non-Euclidean data, and recent research on point cloud semantic and part segmentation show their power in encoding local and global information [31, 15, 14, 29]. In this paper, we use GCNs to design a novel and versatile point upsampling module called NodeShuffle (Figure 2), which is better equipped at encoding local point information and at learning to generate new points instead of merely replicating parts of the input.

Point clouds often represent objects of variable part sizes. Using multi-scale features is an effective way to encode this property and is essential for obtaining point clouds of high quality. Recent works like PU-Net [39] extract point features at different downsampled levels. While such an architecture can encode multi-scale features, down-sampling leads to loss of fine-grained details. In contrast, 3PU [37] proposes a progressive upsampling network using different numbers of neighbors in subsequent upsampling units. This achieves different receptive fields and encodes multi-scale information. However, 3PU is computationally expensive due to its progressive nature. In this paper, we tackle this multi-scale feature learning problem using a multi-path densely connected GCN architecture called Inception DenseGCN. Following its prevalent usage in image recognition [26, 27, 25] for the merits of efficient extraction of multi-scale image information, we adopt the Inception architecture to encode multi-scale point features, after it is modified to use densely connected GCNs instead of CNNs.

Contributions. We summarize our contributions as three-fold. **(1)** We propose *NodeShuffle*, a novel point cloud upsampling module using graph convolutions. We show how NodeShuffle can be seamlessly integrated into current point upsampling pipelines and consistently improve their performance. **(2)** We design *Inception DenseGCN*, a feature extraction block that effectively encodes multi-scale information. We combine Inception DenseGCN and NodeShuffle into a new architecture called *PU-GCN*. As compared to the state-of-the-art, PU-GCN achieves better upsampling quality, requires less parameters, and runs faster. Through extensive quantitative and qualitative experiments and for both synthetic and real data, we show the superior performance of PU-GCN. **(3)** We compile *PU1K*, a new large-scale point cloud upsampling dataset with various levels of shape diversity. We show the challenge of PU1K to current learning-based methods.

2. Related Work

Learning-based point cloud upsampling. Deep learning methods illustrate a promising improvement over their optimization-based counterparts [1, 18, 10, 33] due to their data-driven nature and the learning capacity of neural net-

works. Learning features directly from point clouds was made possible by deep neural networks, such as PointNet [6], PointNet++ [21], DGCNN [31], KPConv [29], etc. Yu *et al.* [39] introduced PU-Net, which learns multi-scale features per point and expands the point set via multi-branch MLPs. However, PU-Net needs to downsample the input first to learn multi-scale features, which causes unnecessary resolution loss. Yu *et al.* [38] also proposed EC-Net, an edge-aware network for point set consolidation. It uses an edge-aware joint loss to encourage the network to learn to consolidate points for edges. However, EC-Net requires a very expensive edge-notation for training. Wang *et al.* [37] proposed 3PU, a progressive network that duplicates the input point patches over multiple steps. 3PU is computationally expensive due to its progressive nature, and it requires more data to supervise the middle stage outputs of the network. Recently, Li *et al.* [16] proposed PU-GAN, a Generative Adversarial Network (GAN) designed to learn upsampled point distributions. While the major contribution and the performance gain is from the discriminator part, the generator architecture receives less attention in their work. Recently, PUGeo-Net[22] proposes to upsample points by learning the first and second fundamental forms of the local geometry. However, their method needs additional supervision in the form of normals, which many point clouds like those generated by LIDAR sensors do not come with. Our work tackles upsampling by leveraging a new Inception based module to extract multi-scale information, and by using a novel GCN-based upsampling module to capture local point information. This avoids the need for additional annotations (*e.g.* edges, normals, point clouds at intermediate resolutions, *etc.*) or a sophisticated discriminator.

3D shape completion. 3D shape completion is intertwined with point cloud upsampling. Researching methods in shape completion can inspire the works of upsampling, and vice versa. Earlier and some recent works on shape completion are based on voxel representations [34, 3, 9, 35] and implicit representations [24]. Point cloud upsampling can also be tackled with these representations, which could be an interesting research direction. More related to point cloud upsampling, point-based completion methods [40, 36, 28, 19, 32] that directly process on the point clouds (point coordinates and the attributes) have recently shown comparable or even superior performance to those voxel or implicit representation based counterparts.

Graph convolutional networks (GCNs). To cope with the increasing amount of non-Euclidean data in real-world scenarios, a surge of graph convolutional networks [13, 8, 30, 20, 31] have been proposed in recent years. Kipf *et al.* [13] simplify spectral graph convolutions with a first-order approximation. DGCNN [31] propose EdgeConv to conduct dynamic graph convolution on point clouds. DeepGCNs [15, 14] introduce residual/dense connections and dilated convolutions to GCNs, and successfully trained deep GCN

architectures. Previous GCN works mainly investigate basic modules of discriminative models. However, due to the unordered and irregular nature of graph data, generative tasks for this modality remain elusive. Recently, Graph U-Nets [5] propose the graph unpooling operation, which can only be used to restore the graph to its original structure and is not able to upsample nodes by an arbitrary ratio. As a focus in this paper, the point upsampling technique, which is an indispensable component for generative models, is under-explored in the GCN domain. We propose a GCN-based upsampling module to tackle the problem.

Multi-scale feature extraction. Inception architectures [26, 27, 25] enable superior performance in image recognition at relatively low computational cost. They extract multi-scale information by using different kernel sizes in different paths of the architecture. Inspired by the success of the Inception architecture for CNNs, Kazi *et al.* [11] proposed InceptionGCN, in which feature maps are passed into multiple branches, then each branch applies one graph convolution with a different kernel size. The outputs of these branches are aggregated by concatenation. We also adopt the Inception concept in our work to propose Inception DenseGCN, a GCN architecture that improves upon InceptionGCN by leveraging dilated graph convolutions, skip connections, and global pooling.

3. Methodology

We propose a novel GCN-based upsampling module (called NodeShuffle) and multi-scale feature extractor (called Inception DenseGCN). We combine Inception DenseGCN and Nodeshuffle into the proposed point cloud upsampling pipeline: PU-GCN.

3.1. Upsampling Module: NodeShuffle

Inspired by PixelShuffle [23] from the image super-resolution literature, we propose *NodeShuffle* to effectively upsample point clouds. NodeShuffle is a graph convolutional upsampling layer, illustrated in Figure 2. Given node features \mathcal{V}_l with shape $N \times C$, the NodeShuffle operator will output the new node features \mathcal{V}_{l+1} with shape $rN \times C$ as follows (r is the upsampling ratio):

$$\mathcal{V}_{l+1} = \mathcal{PS}(\mathcal{W}_{l+1} * \mathcal{V}_l + b_{l+1}), \quad (1)$$

where \mathcal{PS} is a periodic shuffling operator that rearranges the graph (*e.g.* point features) of shape $N \times rC$ to $rN \times C$. The NodeShuffle operation can be divided into two steps. (1) Channel expansion: we use a 1 layer GCN to expand node features \mathcal{V}_l to shape $N \times rC$ using learnable parameters \mathcal{W}_{l+1} and b_{l+1} . (2) Periodic shuffling: we rearrange the output of channel expansion to shape $rN \times C$.

In contrast to multi-branch MLPs [39] or duplicate-based upsampling [37, 16], NodeShuffle leverages graph convolutions instead of CNNs. Although GCNs are common modules for feature extraction, to the best of our knowledge, we

are the first to introduce a GCN-based upsampling module. Our GCN design choice stems from the fact that GCNs enable our upsampler to encode spatial information from point neighborhoods and learn new points from the latent space rather than simply duplicating the original points (as done in 3PU [37] and PU-GAN [16]) or copying points after different transforms through multi-branch MLPs (as done in PU-Net[39]). A more detailed comparison of these upsampling modules can be found in the **supplement**.

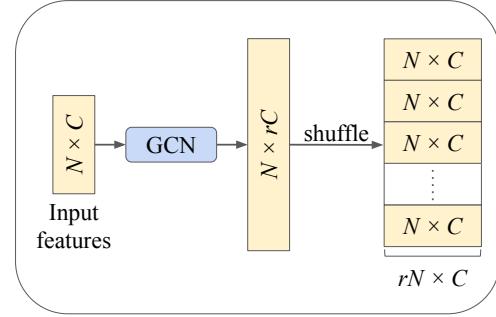


Figure 2: **Upsampling module: NodeShuffle.** We expand the number of input features by r times using a GCN layer, and then apply a shuffle operation to rearrange the feature map. r denotes the upsampling ratio.

3.2. Multi-scale Features: Inception DenseGCN

Point clouds scanned using 3D sensors often include objects of various sizes and point resolutions. In order to encode the multi-scale nature of point clouds, we propose a new *Inception DenseGCN* feature extractor, which effectively integrates the densely connected GCN (DenseGCN) module from DeepGCNs [15] into the Inception module from GoogLeNet [26]. Residual and dense connections have proven to be useful at increasing point cloud processing performance [15]. We favor dense over residual connections here, since the former utilizes features from previous layers, as well as different inception paths.

Figure 3 illustrates our Inception DenseGCN block. The input features are compressed by a bottleneck layer (single layer MLP) at first to reduce the computation in subsequent layers. The compressed features are passed into two parallel DenseGCN blocks. Each DenseGCN block is composed of three layers of densely connected dilated graph convolutions (introduced in DeepGCNs [15]). DenseGCN is defined by the number of node neighbors k (kernel size for GCNs), dilation rate d , and growth rate c . The two DenseGCN blocks inside the Inception DenseGCN have the same kernel size (20) but different dilation rate (1, and 2, respectively) to gain different respective fields without increasing the kernel sizes and FLOPs. Similar to dilated convolutions in the 2D case, the dilated graph convolution efficiently increases the receptive field using the same kernel size without reducing spatial resolution. Additionally,

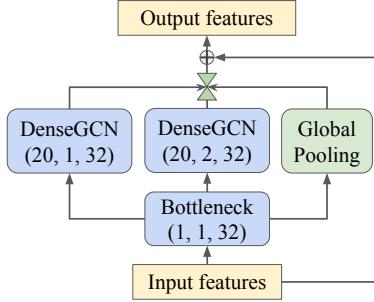


Figure 3: **Multi-scale feature extractor: Inception DenseGCN.** We use the parameters (k, d, c) to define a DenseGCN block. k is the number of neighbors (kernel size), d is the dilation rate, and c is the number of output channels. Green \bowtie denotes feature-wise concatenation. Note that the d is different in the different DenseGCN blocks to achieve the goal of multi-scale feature extraction.

we add a global pooling layer to extract global contextual information. The DenseGCN blocks and the global pooling layer target different receptive fields and therefore allow the Inception module to extract multi-scale information. Each Inception block outputs a concatenation of the DenseGCN blocks and the global pooling layer in addition to the input features. KNN is used to build the graph, whose nodes are the points and edges define the K nearest point neighborhood. Note that KNN is only computed once at the first layer of the Inception DenseGCN block. Experiments show that making the graph dynamic in the later layers inside an Inception DenseGCN module does not actually impact performance much, but it does increase the model’s computational footprint. So, the graph structure is designed to be shared in subsequent DenseGCN layers inside Inception DenseGCN.

The InceptionGCN proposed by Kazi *et al.* [11] is as simple as concatenating multiple GCN layers with different kernel sizes. In contrast, our Inception DenseGCN leverages DenseGCN, global pooling, and skip connections to improve the performance. We further use bottleneck layers and dilated graph convolutions to reduce the computational burden.

3.3. PU-GCN Architecture

We combine Inception DenseGCN, NodeShuffle, and a standard coordinate reconstructor into a new upsampling pipeline called PU-GCN (Figure 4). Given a point cloud of size $N \times 3$, PU-GCN computes point features of size $N \times C$ using the Inception DenseGCN feature extractor. Then, the upsampler transforms the $N \times C$ features to $rN \times C'$ using NodeShuffle. Finally, the coordinate reconstructor generates the $rN \times 3$ upsampled point cloud. We use EdgeConv (proposed in DGCNN [31]) as the default GCN layer.

Inception feature extractor. We use 1 GCN layer at the be-

ginning of PU-GCN to embed the 3D coordinates into latent space. The point embeddings are passed through several Inception DenseGCN blocks. We use two such blocks by default in PU-GCN, and we experiment with the number of Inception DenseGCN blocks in Section 4.6. The outputs of these blocks are combined with residual connections. The final output is passed to our NodeShuffle upsampler.

Upsampler. Our upsampler consists of three components: a bottleneck layer, an upsampling module, and a feature compression layer. Given the input features, we first shrink the input to $N \times C$ by a bottleneck layer (MLP) so as to reduce the computation. Then, we use the proposed NodeShuffle to generate denser features of size $rN \times C$. Finally, we use two sets of MLPs to compress the features to $rN \times C'$.

Coordinate reconstructor. We reconstruct points from latent space to coordinate space, resulting in the desired denser point cloud of size $rN \times 3$. We use the same coordinate reconstruction approach as PU-GAN [16], in which 3D coordinates are regressed using two sets of MLPs.

4. Experiments

We compile a large-scale dataset for point cloud upsampling called PU1K. Quantitative and qualitative results on PU-GAN[16]’s dataset as well as PU1K show the superior performance of PU-GCN. We also conduct extensive ablation studies to show the benefits of the proposed Inception DenseGCN and NodeShuffle upsampling modules.

4.1. Datasets

We propose a new dataset for point cloud upsampling, denoted as *PU1K*. It is nearly 8 times larger than the largest publicly available dataset collected by PU-GAN [16]. PU1K consists of 1,147 3D models split into 1020 training samples and 127 testing samples. The training set contains 120 3D models compiled from PU-GAN’s dataset [16], in addition to 900 different models collected from ShapeNetCore [2]. The testing set contains 27 models from PU-GAN and 100 more models from ShapeNetCore. The models from ShapeNetCore are chosen from 50 different categories. We randomly choose 200 models from each category to obtain 1,000 different models with various shape complexities to encourage diversity. Overall, PU1K covers a large semantic range of 3D objects and includes simple, as well as complex shapes. To show the value of our proposed dataset, we compare our PU-GCN with previous approaches on both PU1K and the latest dataset proposed by PU-GAN, which contains only 120 3D models for training and 27 models for testing.

For training and testing and following common practice, we use Poisson disk sampling from the original meshes to generate pairs of input and ground truth point clouds. For training, we crop 50 patches from each 3D model as the input to the network. In total, we obtain 51,000 training patches in PU1K. Each patch consists of 256 points

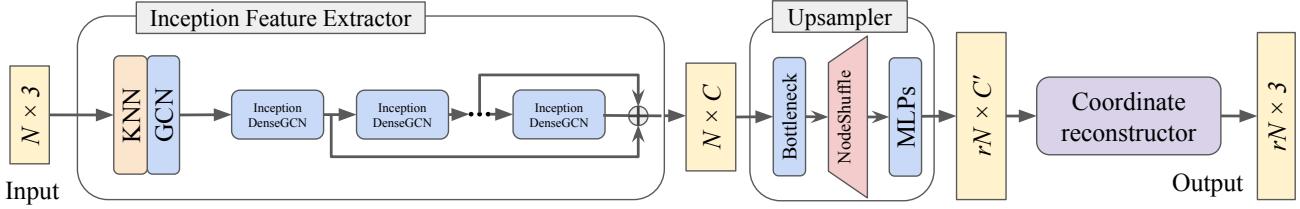


Figure 4: **PU-GCN architecture.** PU-GCN uses an inception feature extractor consisting of one or more Inception DenseGCN blocks, followed by the NodeShuffle based upsampler, and a coordinate reconstructor.

as low resolution input and 1024 points as ground truth. As for testing data, we generate pairs of input point cloud (2048 points) and ground truth (8096 points). In testing, we use farthest point sampling at first to sample overlapping patches (patch size is 256) of the input point cloud and ensure coverage of the entire input. The final result is obtained by first merging the overlapping patch outputs and then re-sampling with farthest point sampling. We will open-source the original meshes of PU1K and the sampled point clouds to standardize the dataset. More details of PU1K and a comparison with PU-GAN’s dataset are in [supplement](#).

4.2. Loss Function and Evaluation Metrics

Loss Function. We use the Chamfer distance loss to minimize the distance of the predicted point cloud and the referenced ground truth in our experiments:

$$C(P, Q) = \frac{1}{|P|} \sum_{p \in P} \min_{q \in Q} \|p - q\|_2^2 + \frac{1}{|Q|} \sum_{q \in Q} \min_{p \in P} \|p - q\|_2^2 \quad (2)$$

where P is the predicted point cloud, Q is the ground truth, p is a 3D point from P , and q is a 3D point from Q . The operator $\|\cdot\|_2^2$ denotes the squared Euclidean norm.

Evaluation Metrics. Following previous work, we use the Chamfer distance (CD), Hausdorff distance (HD), and point-to-surface distance (P2F) w.r.t ground truth meshes as evaluation metrics. The smaller the metrics, the better the performance. We also report the parameter size (Params.) and the inference time. The inference time is reported as the average inference time (over the whole test set in 5 runs) for a model processing one patch containing 256 points. All models are tested on the same computer with one NVIDIA TITAN 2080Ti GPU and an Intel Xeon E5-2680 CPU.

4.3. Implementation Details

We train PU-GCN for 100 epochs with batch size 64 on an NVIDIA TITAN 2080Ti in all the experiments. We optimize using Adam with a learning rate of 0.001 and beta 0.9. Similar to previous work, we perform point cloud normalization and augmentation (rotation, scaling, and random perturbations). We train PU-Net [39], 3PU [37], and our PU-GCN on both the PU1K dataset and PU-GAN’s dataset. Here, we note that we are unable to reproduce PU-GAN’s

Table 1: **Comparison of PU-GCN vs. state-of-the-art on PU-GAN’s dataset.** PU-GCN outperforms PU-Net, 3PU, and PU-GAN on nearly all metrics with the least parameters and fastest inference. **Bold** denotes the best performance.

Network	CD _{10⁻³}	HD _{10⁻³}	P2F _{10⁻³}	Param. Kb	Time ms
PU-Net [39]	0.556	4.750	4.678	814.3	10.04
3PU [37]	0.298	4.700	2.855	76.2	10.86
PU-GAN [16]	0.280	4.640	2.330	684.2	14.28
PU-GCN	0.258	1.885	2.721	76.0	8.83

[16] results from the code made available by its authors, most probably because of the unstable nature of the inherent GAN architecture in PU-GAN. Therefore, we only compare with PU-GAN on PU-GAN’s dataset using their provided pre-trained model. The very recent work PUGeo-Net[22] is not included in the comparison for the following reasons: (1) their code and pre-trained model are not available; (2) they only conduct experiments on their own dataset that has not been released; (3) they need additional supervision in the form of normals, which are not directly accessible in point clouds. As suggested by previous methods, we use the model from the last epoch to perform the evaluation. We report results using a $\times 4$ upsampling rate, *i.e.* $r = 4$. We did not experiment with a large upsampling factor like $r = 16$, but one can simply apply our pretrain models twice.

4.4. Quantitative and Qualitative Results

Quantitative results on PU-GAN’s dataset. Table 1 reports the performance of our PU-GCN compared to PU-Net[39], 3PU[37], and PU-GAN[16] on PU-GAN’s dataset. PU-GCN maintains significant improvement over 3PU and PU-Net in all metrics, showing the importance of the Inception DenseGCN feature extractor and the NodeShuffle upsampling module. Although PU-GAN leverages an adversarial loss for performance gains, we also outperform PU-GAN in terms of CD and HD, without the need for a GAN architecture. It is also important to mention that PU-GCN is the most parameter-saving and the most efficient architecture among all the models. Compared to PU-GAN, our PU-GCN uses only about 10% of the parameters and speeds

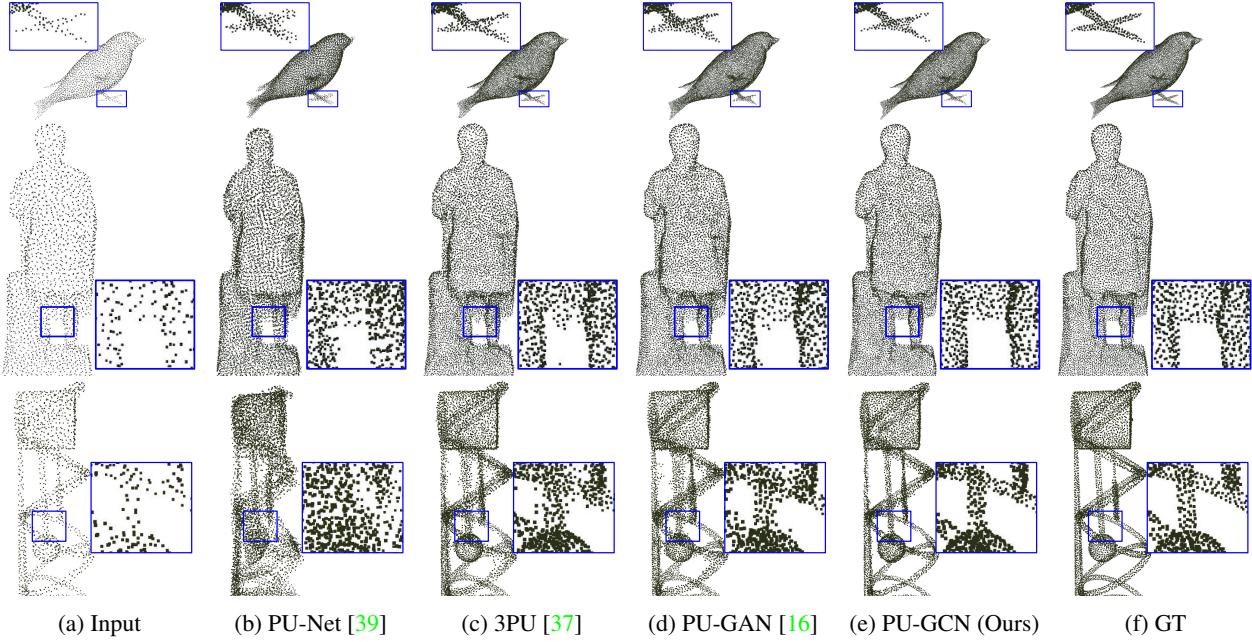


Figure 5: **Qualitative upsampling results.** We show the $\times 4$ upsampled results of input point clouds (2048 points) when processed by different upsampling methods (PU-Net, 3PU, PU-GAN and our PU-GCN). PU-GCN produces the best results overall, while generating less outliers and preserving fine-grained local details (refer to close-ups).

up inference time by more than 40%. For the P2F metric, we do not outperform PU-GAN quantitatively. However, the qualitative results in Figure 5 show that PU-GCN generates fewer outliers and higher quality local and fine-grained details (*e.g.* the legs of the bird in row 1) than PU-GAN in general, even though these upsampled point clouds have higher P2F values than those of PU-GAN. This discrepancy indicates that P2F might not be as reliable a metric as CD and HD for point cloud upsampling, especially since the latter distances are meant for point clouds (the natural form of the input/output in training) and not meshes (as is the case for P2F). More examples and an analysis of this discrepancy can be found in the **supplement**.

Quantitative results on the PU1K dataset. Table 2 compares PU-GCN against PU-Net and 3PU on PU1K. We do not compare against PU-GAN on this dataset for the reason mentioned in Section 4.3. We observe that all methods achieve lower performance overall (especially in CD and HD) when trained and evaluated on PU1K as compared to PU-GAN’s dataset. As such, PU1K presents a challenge to state-of-the-art methods compared to the much smaller and less diverse PU-GAN’s dataset. PU-GCN also clearly outperforms PU-Net and 3PU in all three metrics on this new challenging dataset.

Qualitative results. Figure 5 shows qualitative upsampling results generated by PU-GCN and the state-of-the-art methods. We note that all models used here are trained on PU-GAN’s dataset for a fair comparison with PU-GAN. Up-

Table 2: **Comparison of PU-GCN vs. state-of-the-art on PU1K.** PU-GCN outperforms PU-Net and 3PU. We do not compare against PU-GAN on this dataset, since we are unable to reproduce PU-GAN results from the publicly available code. **Bold** denotes the best performance.

Network	CD \downarrow 10^{-3}	HD \downarrow 10^{-3}	P2F \downarrow 10^{-3}	Param. Kb	Time ms
PU-Net[39]	1.155	15.170	4.834	814.3	10.04
3PU[37]	0.935	13.327	3.551	76.2	10.86
PU-GCN	0.585	7.577	2.499	76.0	8.83

sampled point clouds and their close-ups show that PU-GCN produces fewer outliers, while preserving more fine-grained details. Specifically, close-ups of the bird point cloud (top row) show that PU-GCN successfully upsamples intricate structures of input points. In addition, the statue (second row) shows how PU-GCN succeeds in upsampling with much fewer outliers. We also observe that other methods tend to merge originally separate structures as shown in the example of the clock (third row), while our method preserves this separation with high quality. The qualitative results clearly show the effectiveness of our multi-scale feature extractor (Inception DenseGCN) and our GCN based upsampling module (NodeShuffle) in capturing detailed local information. More qualitative results (especially for models trained on PU1K dataset) and a discussion of some failure cases can be found in the **supplement**.

4.5. Upsampling Real-Scanned Point Clouds

Figure 6 compares PU-GCN against its two most competitive upsampling methods (3PU and PU-GAN) on real-scanned data from the KITTI dataset [7]. KITTI contains challenging real-world scenes and objects. All the models are trained on PU-GAN’s dataset for a fair comparison with PU-GAN. To ease visualization and since it consistently produces worse results, PU-Net [39] is not included here. We observe that the other methods tend to produce more outliers and overfill holes (*e.g.* the first close up on the window of the car), while PU-GCN encodes local information well and preserves these fine-grained details. Our method also does very well when upsampling objects of interest, such as motorcycle as shown in the second close up. While other methods tend to merge the pedal and body of the motorcycle and ruin its original shape, PU-GCN successfully produces a higher level of detail and structure. More examples of real-scanned point clouds from different datasets can be found in the **supplement**.

4.6. Ablation Study

We conduct ablation studies on NodeShuffle and Inception DenseGCN. The baseline model is PU-GCN equipped with two Inception DenseGCN and the NodeShuffle upsampler. All the models are trained and evaluated on PU1K.

Inception Modules. We validate the effectiveness of our Inception DenseGCN by replacing it with a DenseGCN. The results in Table 3 show that our Inception DenseGCN outperforms the DenseGCN in all metrics. The second row demonstrates that the Inception DenseGCN with multiple receptive fields achieved by using two DenseGCN blocks improves the upsampling quality a lot with similar latency, compared to Inception DenseGCN with only a single DenseGCN block. Additionally, using DenseGCN in Inception DenseGCN works better by a large margin than using GCN (Inception GCN) as expected (third row). The fourth row shows that using dilated graph convolutions can achieve better performance, while being more efficient compared to the Inception module using different kernel sizes. We further show that using residual connections and global pooling inside Inception DenseGCN improves performance with a negligible effect on inference speed (row 5 and row 6). We also study the effect of the number of Inception DenseGCN blocks. PU-GCN with two Inception blocks outperforms PU-GCN with only one Inception block, thus showing that inserting one more Inception DenseGCN block can further improve upsampling performance. We also experimented with using more blocks, but performance did not improve while computational complexity did. Therefore, we use two Inception DenseGCN blocks in PU-GCN by default.

Upsampling Modules. We show the effectiveness of our NodeShuffle by integrating it into different upsampling architectures and replacing the original upsampling modules.

Table 3: **Ablation study on Inception DenseGCN.** Inception DenseGCN performs better than a DenseGCN. Using two DenseGCN blocks in Inception DenseGCN to extract multi-scale information is better than using only one block. As expected, using DenseGCN in Inception DenseGCN works better than using GCN (Inception GCN). Using dilated graph convolution instead of regular graph convolution with different kernel sizes also achieves better performance with faster inference. Residual connections and global pooling inside Inception DenseGCN further improve performance. Increasing the number of Inception DenseGCN blocks tends to improve PU-GCN performance.

Ablation	CD _{10⁻³}	HD _{10⁻³}	P2F _{10⁻³}	Param. Kb	Time ms
DenseGCN (w/o Inception)	0.753	10.691	3.103	56.16	9.89
single DenseGCN block	0.630	9.428	2.608	56.32	8.82
Inception GCN	0.675	9.951	2.723	59.59	8.65
w/o dilated graph convolution	0.624	8.871	2.530	75.97	8.91
w/o residual connection	0.621	8.979	2.603	75.97	8.74
w/o global pooling	0.663	9.875	2.665	75.88	8.74
1 Inception DenseGCN	0.639	9.051	2.582	59.33	6.40
PU-GCN	0.585	7.577	2.499	75.97	8.83

Results in Table 4 clearly show that NodeShuffle helps both PU-Net and 3PU reach better performance with less parameters and negligible computational overhead ($\leq 1\text{ms}$ latency). We also study the effectiveness of GCN inside NodeShuffle by replacing it with a set of MLPs, called MLPShuffle. GCN outperforms the MLPs counterpart. Table 4 also shows that the proposed Inception DenseGCN outperforms the feature extractors used in PU-Net and 3PU, when comparing PU-GCN (NodeShuffle) with PU-Net (NodeShuffle) and 3PU (NodeShuffle). A qualitative comparison of the original 3PU, 3PU (NodeShuffle), and PU-GCN is shown in Figure 1. NodeShuffle shows a clear improvement in generating samples with less noise and better details, as compared to the original upsampling method used in 3PU. Inception DenseGCN further improves the performance by preserving better intricate structures.

Table 4: **Ablation study on NodeShuffle.** Results show that NodeShuffle can transfer well to different upsampling architectures in the literature. Replacing the original upsampling module with NodeShuffle improves upsampling performance overall. The effectiveness of the GCN layer in NodeShuffle is shown when replacing the GCN layer with a set of MLPs (MLPShuffle).

Network	CD 10 ⁻³	HD 10 ⁻³	P2F 10 ⁻³	Param. Kb	Time ms
PU-Net (Original) [39]	1.155	15.170	4.834	814.3	10.04
PU-Net (NodeShuffle)	0.974	13.522	4.474	462.1	11.04
3PU (Original) [37]	0.935	13.327	3.551	76.2	10.86
3PU (NodeShuffle)	0.780	10.462	3.228	71.4	11.78
PU-GCN (Duplicate)	0.788	11.269	3.031	74.2	8.63
PU-GCN (MLPShuffle)	0.682	10.692	2.586	76.2	8.87
PU-GCN (NodeShuffle)	0.585	7.577	2.499	76.0	8.83

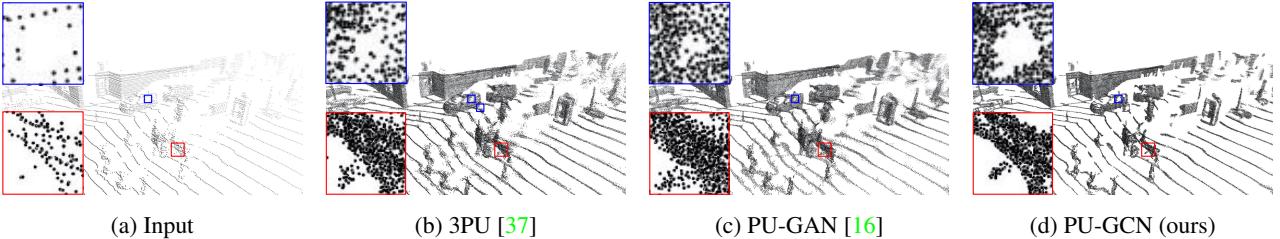


Figure 6: **Upsampling real-scanned point clouds from KITTI [7]**. PU-GCN preserves intricate structures and generates fine-grained details (*e.g.* the window of the car, the pedal of the motorcycle, *etc.*). Please zoom in to see details.

4.7. Effects of Additive Noise and Input Sizes

Upsampling noisy point clouds. To show the robustness of PU-GCN, we perturb the input point cloud with additive Gaussian noise at varying noise levels. We only compare PU-GCN against PU-GAN[16] since PU-GAN is the state-of-the-art. Both models are trained using the same augmentation strategy of point cloud perturbation. Qualitative results in Figure 7 show that PU-GCN can preserve fine-grained details with very few outliers, even in the presence of additive noise, while PU-GAN tends to produce more outliers.

Upsampling point clouds of varying sizes. Figure 8 shows qualitative examples of upsampling point clouds with PU-GCN for different input sizes. Our PU-GCN always produces high quality point clouds for the range of input sizes. This indicates that even if PU-GCN is trained on patches with only 256 points, it can generalize to point clouds with different sizes. As expected, PU-GCN generates better quality results when the input point cloud is denser.

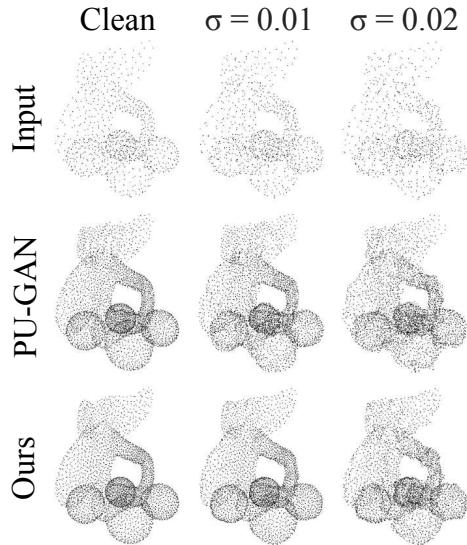


Figure 7: **Effect of additive noise.** The top row shows the input point clouds with different additive noise levels. The middle and bottom rows show the point clouds upsampled by PU-GAN[16] and our PU-GCN, respectively. PU-GCN preserves more fine-grained details with fewer outliers.

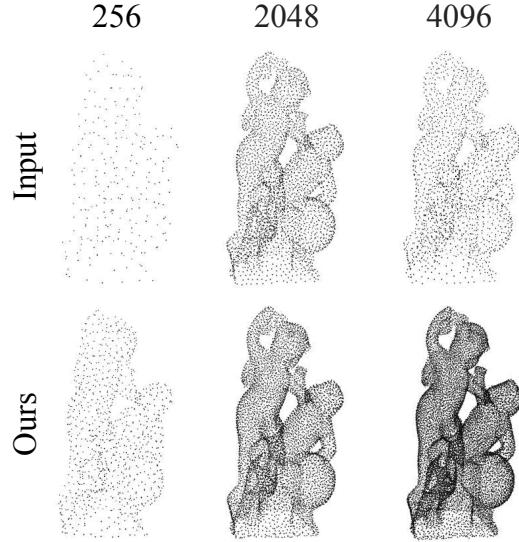


Figure 8: **Effect of input size.** The top row shows the input point clouds of different sizes, and the bottom row shows the $\times 4$ upsampled outputs of our PU-GCN. PU-GCN always produces high quality results regardless of input size.

5. Conclusion

We propose a novel GCN based point cloud upsampling module called NodeShuffle, which improve state-of-the-art upsampling pipelines when it is used in place of the original upsampling. We also introduce the Inception DenseGCN to encode multi-scale information. We further compile and introduce a new large-scale dataset PU1K for point cloud upsampling. Extensive experiments demonstrate that our proposed PU-GCN pipeline, which integrates Inception DenseGCN and NodeShuffle, outperforms state-of-the-art methods on PU1K and another dataset, while requiring fewer parameters and being more efficient in inference. We also show that PU-GCN produces higher upsampling quality on real-scanned point clouds compared to other methods.

Acknowledgments. The authors thank Silvio Giancola and Chen Zhao for their suggestive comments. This work was supported by KAUST Office of Sponsored Research through the Visual Computing Center (VCC) funding.

References

- [1] Marc Alexa, Johannes Behr, Daniel Cohen-Or, Shachar Fleishman, David Levin, and Cláudio T. Silva. Computing and rendering point set surfaces. *IEEE Trans. Vis. Comput. Graph.*, 9:3–15, 2003. 2
- [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. 4
- [3] Angela Dai, C. Qi, and M. Nießner. Shape completion using 3d-encoder-predictor cnns and shape synthesis. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6545–6554, 2017. 2
- [4] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38:295–307, 2014. 1
- [5] H. Gao and S. Ji. Graph u-nets. In *ICML*, 2019. 3
- [6] Alberto Garcia-Garcia, Francisco Gomez-Donoso, J. A. Rodriguez, Sergio Orts, Miguel Cazorla, and Jorge Azorín López. Pointnet: A 3d convolutional neural network for real-time object class recognition. *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 1578–1584, 2016. 2
- [7] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012. 7, 8
- [8] Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems*, pages 1024–1034, 2017. 2
- [9] Xiaoguang Han, Z. Li, Haibin Huang, E. Kalogerakis, and Y. Yu. High-resolution shape completion using deep neural networks for global structure and local geometry inference. *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 85–93, 2017. 2
- [10] Hui Huang, Shihao Wu, Minglun Gong, Daniel Cohen-Or, Uri M. Ascher, and Hao Zhang. Edge-aware point set resampling. *ACM Trans. Graph.*, 32:9:1–9:12, 2013. 2
- [11] Anees Kazi, Shayan Shekarforoush, S Arvind Krishna, Hendrik Burwinkel, Gerome Vivar, Karsten Kortüm, Seyed-Ahmad Ahmadi, Shadi Albarqouni, and Nassir Navab. Inceptiongcn: Receptive field aware graph convolutional network for disease prediction. In *International Conference on Information Processing in Medical Imaging*, pages 73–85. Springer, 2019. 3, 4
- [12] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1646–1654, 2015. 1
- [13] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016. 2
- [14] Guohao Li, Matthias Müller, Guocheng Qian, Itzel C. Delgadillo, Abdulellah Abualshour, Ali K. Thabet, and Bernard Ghanem. Deepgcns: Making gcns go as deep as cnns. *ArXiv*, abs/1910.06849, 2019. 2
- [15] Guohao Li, Matthias Müller, Ali Thabet, and Bernard Ghanem. Deepgcns: Can gcns go as deep as cnns? In *The IEEE International Conference on Computer Vision (ICCV)*, 2019. 2, 3
- [16] Ruihui Li, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. Pu-gan: a point cloud upsampling adversarial network. In *IEEE International Conference on Computer Vision (ICCV)*, 2019. 1, 2, 3, 4, 5, 6, 8
- [17] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1132–1140, 2017. 1
- [18] Yaron Lipman, Daniel Cohen-Or, David Levin, and Hillel Tal-Ezer. Parameterization-free projection for geometry reconstruction. In *ACM SIGGRAPH 2007 Papers, SIGGRAPH ’07*, New York, NY, USA, 2007. ACM. 2
- [19] Minghua Liu, Lu Sheng, Shilin Yang, J. Shao, and Shi-Min Hu. Morphing and sampling network for dense point cloud completion. *ArXiv*, abs/1912.00280, 2020. 2
- [20] Trang Pham, Truyen Tran, Dinh Phung, and Svetha Venkatesh. Column networks for collective classification. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017. 2
- [21] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J. Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *NIPS*, 2017. 2
- [22] Yue Qian, J. Hou, S. Kwong, and Y. He. Pugeo-net: A geometry-centric network for 3d point cloud upsampling. *ArXiv*, abs/2002.10277, 2020. 2, 5
- [23] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P. Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1874–1883, 2016. 1, 3
- [24] David Stutz and Andreas Geiger. Learning 3d shape completion from laser scan data with weak supervision. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1955–1964, 2018. 2
- [25] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alex Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *AAAI*, 2016. 2, 3
- [26] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott E. Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, 2014. 2, 3
- [27] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, 2015. 2, 3
- [28] Lyne P. Tchapmi, V. Kosaraju, Hamid Rezatofighi, I. Reid, and S. Savarese. Topnet: Structural point cloud decoder. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 383–392, 2019. 2

- [29] Hugues Thomas, Charles R. Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, François Goulette, and Leonidas J. Guibas. Kpconv: Flexible and deformable convolution for point clouds. *Proceedings of the IEEE International Conference on Computer Vision*, 2019. [2](#)
- [30] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017. [2](#)
- [31] 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, 4](#)
- [32] Yida Wang, D. J. Tan, N. Navab, and Federico Tombari. Soft-poolnet: Shape descriptor for point cloud completion and classification. *ArXiv*, abs/2008.07358, 2020. [2](#)
- [33] Shihao Wu, Hui Huang, Minglun Gong, Matthias Zwicker, and Daniel Cohen-Or. Deep points consolidation. *ACM Trans. Graph.*, 34(6):176:1–176:13, Oct. 2015. [2](#)
- [34] Zhirong Wu, Shuran Song, A. Khosla, F. Yu, Linguang Zhang, X. Tang, and J. Xiao. 3d shapenets: A deep representation for volumetric shapes. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1912–1920, 2015. [2](#)
- [35] Haozhe Xie, Hongxun Yao, Shangchen Zhou, Jiageng Mao, S. Zhang, and Wenxiu Sun. Grnet: Gridding residual network for dense point cloud completion. *ArXiv*, abs/2006.03761, 2020. [2](#)
- [36] Y. Yang, Chen Feng, Y. Shen, and Dong Tian. Foldingnet: Point cloud auto-encoder via deep grid deformation. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 206–215, 2018. [2](#)
- [37] Wang Yifan, Shihao Wu, Hui Huang, Daniel Cohen-Or, and Olga Sorkine-Hornung. Patch-based progressive 3d point set upsampling. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019. [1, 2, 3, 5, 6, 7, 8](#)
- [38] Lequan Yu, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. Ec-net: an edge-aware point set consolidation network. In *ECCV*, 2018. [1, 2](#)
- [39] Lequan Yu, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. Pu-net: Point cloud upsampling network. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. [1, 2, 3, 5, 6, 7](#)
- [40] Wentao Yuan, Tejas Khot, David Held, C. Mertz, and M. Hebert. Pcn: Point completion network. *2018 International Conference on 3D Vision (3DV)*, pages 728–737, 2018. [2](#)