ITC 1/53
Information Technology and Control
Vol. 53 / No. 1 / 2024
pp.160-170
DOI 10.5755/j01.itc.53.1.34047

Point Cloud Completion Based on Nonlocal Neural Networks with Adaptive Sampling

Received 2023/05/06 | Accepted after revision 2023/09/30

HOW TO CITE: Xing, N., Wang, J., Wang, Y., Ning, K., Chen, F. (2024). Point Cloud Completion Based on Nonlocal Neural Networks with Adaptive Sampling. *Information Technology and Control*, 53(1), 160-170. https://doi.org/10.5755/j01.itc.53.1.34047

Point Cloud Completion Based on Nonlocal Neural Networks with Adaptive Sampling

Na Xing, Jun Wang, Yuehai Wang, Keqing Ning, Fuqiang Chen

School of Information, North China University of Technology, No. 5, Jinyuanzhuang Road, Shijingshan District, Beijing, China; e-mails: xingna@ncut.edu.cn; wangjun18838275466@126.com; wangyuehai@ncut.edu.cn, ningkq@ncut.edu.cn, 18041965208@163.com

Corresponding author: xingna@ncut.edu.cn

Raw point clouds are usually sparse and incomplete, inevitably containing outliers or noise from 3D sensors. In this paper, an improved SA-Net based on an encoder-decoder structure is proposed to make it more robust in predicting complete point clouds. The encoder of the original SA-Net network is very sensitive to noise in the feature extraction process. Therefore, we use PointASNL as the encoder, which weights around the initial sampling points through the AS module (Adaptive Sampling Module) and adaptively adjusts the weight of the sampling points to effectively alleviate the bias effect of outliers. In order to fully mine the feature information of point clouds, it captures the neighborhood and long-distance dependencies of sampling points through the LNL module (Local-NonLocal Module), providing more accurate information for point cloud processing. Then, we use the encoder to extract local geometric features of the incomplete point cloud at different resolutions. Then, an attention mechanism is introduced to transfer the extracted features to a decoder. The decoder gradually refines the local features to achieve a more realistic effect. Experiments on the ShapeNet data set show that the improved point cloud completion network achieves the goal and reduces the average chamfer distance by 3.50% compared to SA-Net.

KEYWORDS: Point Cloud Completion, Nonlocal Neural Network, Adaptive Sampling.

1. Introduction

A point cloud is a type of data that describes three dimensional (3D) objects and scenes. Due to the impact of camera resolution, occlusion and other factors, the point cloud collected by the sensor is usually sparse, incomplete and noisy, which cannot describe the real geometric information of the object. This is not enough for many downstream applications, such as object detection, scene understanding, robot navigation, and augmented reality. Therefore, completion the missing regions of 3D shape is a critical task in the field of computer vision.

Existing methods have achieved remarkable results in shape completion with 3D data as input, such as voxel [4, 15, 39], mesh [6, 14, 24], multi-view images [8, 11, 32] and point cloud [13, 19, 20]. Voxel representation generalizes 3D data to 2D grids, and then applies 2D convolutional neural networks to reconstruct object shapes. However, its computational complexity is high, and as the resolution increases, the required memory increases exponentially. Mesh representation is the use of a model to warp a planar image into a target geometry. However, using meshes to represent complex 3D objects is not easy. Multi-view image representation is the rendering of 3D shapes from a series of 2D views from different angles. However, the reconstructed 3D shape is still insufficient to represent the original information. In contrast, point cloud has become the mainstream data form in 3D vision tasks due to its advantages of small memory usage, low computational complexity, and strong ability to restore object characteristics [32]. Since 3D point clouds are irregular and disordered and do not conform to the regular grids in 2D images, convolutional neural networks are difficult to apply to analyze point cloud data.

In recent years, with the continuous improvement of computing performance and the emergence of large-scale complete 3D shape datasets, such as the ShapeNet dataset. Some deep learning-based methods have achieved great success in classifying and segmenting 3D objects. AGNet [10] treats point cloud data as a graph structure, where each point is a node of the graph. It introduces an attention mechanism to better aggregate the characteristics of neighborhood nodes and map them to predefined category labels to achieve segmentation. Pierdicca [18] uses an improved DGCNN for semantic segmentation of complex, highly variable 3D point cloud historical building models to accelerate the identification of historical building elements. Karolis et al. [21] proposed using computer-aided technology for depth video stream masking and training with convolutional neural networks to obtain more accurate human body segmentation results. This framework has good accuracy and real-time performance in human body segmentation. However, due to factors such as noise or occlusion, the 3D model collected by the sensor is missing, which seriously affects subsequent applications. Since deep learning has shown strong capabilities in some basic applications of 3D objects, and 3D objects are often represented using point cloud data, many point cloud completion methods based on deep learning have been proposed. However, some advanced models result in local details loss, or fail to generate fine shapes, even though the output is a complete structure. PCN (Point Completion Network) utilizes multi-layer perceptrons (MLPs) and max-pooling to refine the point set [33], but it still has unacceptable detail loss. AtlasNet and FoldingNet typically learn global representation from partial point clouds and use global features to generate complete shapes [6, 31], while the completion effect is not satisfactory. TopNet is proposed to use a layered rooted tree structure encoder, which considers the local structural details [23]. However, the output structures of these methods are fuzzy, which is not tolerated by the shape completion. With the development of deep learning, some methods have designed complex neural network models and achieved good shape completion results. They can express the geometric information of point cloud with high-dimensional features, and thus decode the complete shape from the global and local features. The decoder of SA-Net gradually generates refined fine local structural details by folding blocks [25]. PF-Net fuses the extracted global features under different resolutions to generate missing structure by a generator and a discriminator [9]. Some models use hybrid methods to process point cloud data. HFCNN [36] first uses a convolutional neural network to extract features from point cloud data, and then combines global features and local geometric features, using a hybrid method to effectively combine these two features to capture the complex characteristics of point cloud data. Kulikajevas et al. [12] used 3DCNN and 3DGCN to encode and decode different features respectively, and then fused these features together to get a more comprehensive shape representation. The attention mechanism is a very effective feature extraction method that improves the performance and accuracy of the model by focusing on key features in the input data and ignoring irrelevant information. In order to capture effective local geometric structures, SDA-Net [5] embeds the attention mechanism into MLP to extract local deep attention features. AGNet [10] builds a topology structure by creating directed edges from the neighbors of each point, and uses an attention mechanism to select the most important features in the topology structure. Therefore, it is necessary to further explore the local information.

In this paper, we propose an improved SA-Net. To obtain the realistic local structure, we use PointASNL in the encoder, which includes an adaptive sampling (AS) module and a local-nonlocal (LNL) module. The AS module can not only reduce the impact of outliers on feature extraction by adjusting the coordinates of sampling points, but also help to fit a smooth geometric submanifold. The LNL module considers the local information interaction between each sampling point and adjacent regions, as well as the influence of eachsampling point with the entire points in a certain level. It could make full use of input information to benefit learning features.

2. Related Works

In this section, according to the network structure used in point cloud completion, existing architectures can be mainly divided into GAN-based methods, Transformer-based methods and Point-based methods.

2.1. GAN-based Methods

The GAN [3] architecture utilizes implicit learning of the discriminator to refine the set of points provided by the generator. Effective latent space representation of point clouds provides important basic information for 3D shape reconstruction. Chen et al. [2] proposed the use of GAN network to solve the problem of mismatching between completed point cloud and real point cloud data and generate high-resolution 3D shapes. Sarmad et al. [22] proposed Rl-Gan-Net, which captures the latent spatial representation of point clouds through autoencoders, and GANs utilize it to reconstruct missing regions close to real point clouds. Wen et al. [26] emphasized the potential relationship of backward learning from complete shapes to incomplete shapes, using two simultaneous transformation networks, Cycle4Completion, to generate high-fidelity point cloud models. Zhang et al. [39] proposed ShapeInversion, which introduced GAN inversion for the first time to realize shape Status completion. ShapeInversion does not require prior pairing of training data and benefits from pre-training on complete shapes using GANs, querying latent codes to find complete shapes corresponding to missing objects. Due to the characteristics of 3D data, point cloud completion methods based on GAN networks will, on the one hand, produce uneven point clouds and noise during the reconstruction process, and on the other hand, rely heavily on prior knowledge.

2.2. Transformer-based Methods

Transformer was originally introduced as a natural language processing (NLP) framework based on the attention mechanism. It usually adopts an encoder-decoder structure and uses the self-attention mechanism and the cross-attention mechanism to extract features. Due to the self-attention mechanism, Transformer is very effective in handling tasks involving long sequences. In the decoder, the cross-attention mechanism utilizes the information in the encoder to learn attention mappings of query features, making the Transformer powerful in generative tasks. With the emergence of PCT [7], Pointformer [17] and PointTransformer [34], Transformer began to demonstrate powerful learning capabilities in point cloud processing. PMP-Net++ [27] improves PMP-Net [28] and introduces Transforme-enhanced representation learning network, which significantly improves the performance of the point cloud completion network. However, Transformer-based models have some limitations. Due to the large number of Transforme parameters, it cannot be easily deployed on the device compared to other methods.

2.3. Point-based Methods

Since the pioneering of PointNet [19] and PointNet++ [20], point-based methods usually utilize multi-layer perceptrons (MLPs) to capture local features point by point. Due to its simplicity and efficiency, some works use point-wise MLP for point cloud completion.

FoldingNet [31] can reconstruct arbitrary point clouds from a 2D mesh. It applies a virtual force to deform or stretch the 2D mesh onto a 3D surface. This method reconstructs objects with detailed structures and the reconstruction error is small. PCN adopts a coarse-to-fine strategy and uses a multi-layer perceptron to refine local features [33]. Although this method generates a uniform and complete point cloud, the max-pooling aggregation operation leads to information loss during encoding and fails to generate high-fidelity results. Xu et al. [1] designed FinerPCN by combining PCN and point-wise convolution to generate fine complete point clouds with a coarse-to-fine strategy by considering local information. MSN generates fine point clouds by learning point-wise residuals of point clouds [16]. SA-Net uses jumping attention mechanism and multiple folding operations to generate fine-grained point clouds [25]. However, these two methods only represent the global shape, resulting in the loss of local details. NSFA pays special attention to the importance of local features, and will use different scales of local features to construct and refine points in missing regions [35]. We found through research that the above methods are very sensitive to noise or outliers or generate noise during the reconstruction process, resulting in the generated point cloud model not being smooth enough, and the extracted local feature information is not rich enough. In the following sections, our solution will be presented.

3. Our Network Architecture

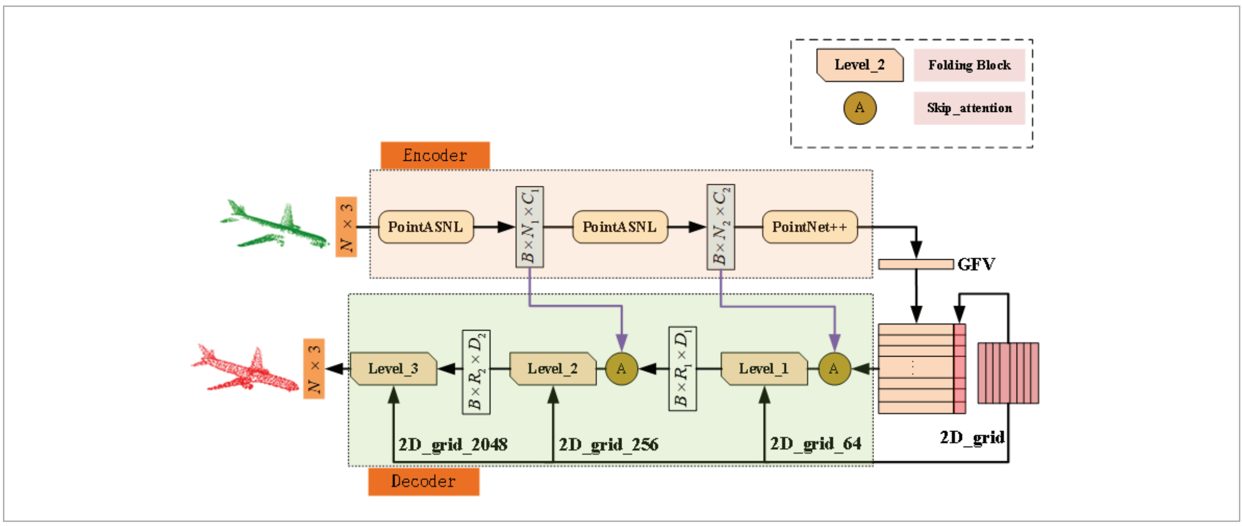
Our network model is shown in Figure 1. It is an encoder-decoder structure. Its detailed design is described below.

3.1. Motivation

Let a point set $Z=\{z_i|i=1,2,3,\ldots\ldots,N\}$ represent a partial 3D object shape by N points, which may contain noise or outliers. The point set of its complete shape is defined as $C=\{c_i|i=1,2,3,\ldots\ldots,N\}$. Please note that although these two sets contain the same number of points, there is no corresponding relationship between geometry Z and C, they just represent the same target shape. Our goal is to design a point cloud completion network model M, input the partial cloud Z into M, and output the reconstructed point cloud model R, where the set of R is expressed as $R=\{r_i \mid i=1,2,3,\cdots\cdots,N\}$, making $C-(C\cap R)=\varnothing$. A high-performance point cloud completion network model should not only complete missing regions, but also be able to generate a detailed local geometric structure. However, most of the existing point cloud completion models are based on encoder-decoder design, and cannot generate high-fidelity output and fine local geometric details, and the richness of features extracted by the encoder directly affects the decoder. The quality of the reconstructed point cloud. We try to generate high-quality point clouds by introducing a new encoder.

Figure 1

Point cloud completion based on nonlocal neural networks with adaptive sampling



3.2. Encoder

We improve the encoder of SA-Net [25], which uses the PointASNL framework as the backbone [30]. The point cloud data with missing information is directly input to the encoder network. The number of points is $N = 2048$, and the position of each point i is (x_i, y_i, z_i). In order to facilitate processing, we use the farthest point sampling (FPS) to obtain the sampling points, where the first level $N_1 = 512$, the second level $N_2 = 256$. The sampling points set is defined as $S = \{s_1, s_2, \ldots, s_i, \ldots, s_m | m = N_1 \text{ or } N_2\}$. Then, the output of the first level is then used directly as the input of the second level to extract global features.

Next, the function of AS module is introduced, which can solve the noise problem. The coordinates and geometric features of the sampling points are input into the AS module for updating. The sampling point s_i can be taken as an example. Firstly, the AS module uses the k-NN algorithm to obtain the k nearest neighbors of s_i, which can be expressed as $\varphi(s_i) = \{s_{i,1}, s_{i,2}, s_{i,3}, \ldots\ldots, s_{i,k}\}$. The corresponding feature set is defined as $F(s_i) = \{f_{i,1}, f_{i,2}, f_{i,3}, \cdots\cdots, f_{i,k}\}$. Then, we use multi-layer perceptions (MLPs) with the Softmax function to obtain the correlation between the sample point s_i and each neighbor point, represented by normalized weights. Finally, coordinates of sampling point $s_i^{\wedge}$ and feature $f_i^{\wedge}$ are updated by the weighted sum of the neighbor points.

As shown in Figure 2, given an airplane, we can see that the input point clouds contain noise. The noise makes the fluid of the airplane appear rough, and the point distribution on the head is scattered. After using the AS module, we can obviously see that the noise points disappear, the outline of the aircraft is smoother, and the overall point distribution is more uniform.

In order to enhance the context learning ability of point clouds and robustness, we combine the most popular point local (PL) cell with point nonlocal (PNL) cell as local-nonlocal (LNL) module. As shown in Figure 3, the k-NN algorithm is used at the sampling points to

Figure 2

The AS module by taking an airplane as an example

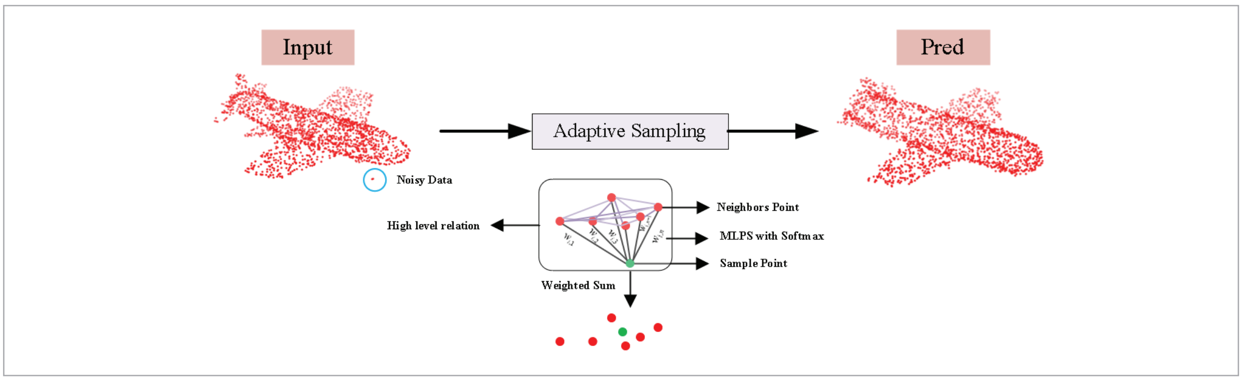
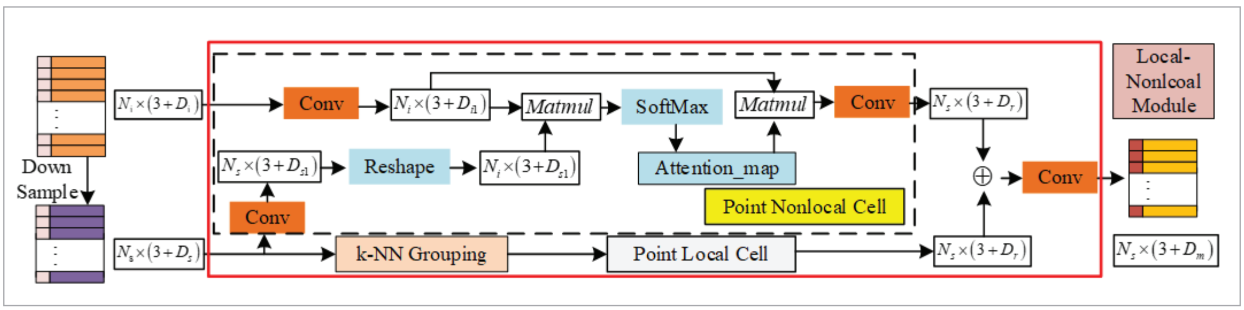


Figure 3

The LNL module. N_s is the number of sampling points. D_i, D_s, ..., D_m are the number of channels. Conv is the convolution operation



obtain multiple groups. The PL cell extracts and aggregates features of each group by MLPs and max-pooling to form local features. Then, multiple local features are aggregated to form global feature. However, PL cell ignores the relationship between the sampling points and the entire points in a certain level. This makes it impossible to get precise information about missing regions. Therefore, we introduce the PNL cell to obtain the context information by fusion. The similarity between sampling points and entire points is obtained by Softmax, normalized weights and entire points weighting. Finally, convolution is used to fuse local and global information. In addition, the information loss in the down-sampling process is reduced with the increase of the number of output channels.

3.3. Decoder

Many decoders use MLP and FC to complete missing regions, but the geometric details are very rough. The reason is that these decoders do not take geometric differences into account. In this letter, we solve these problems with folding blocks, whose backbone framework is up-down-up.

We expand some points in the missing regions. The folding block uses the up-module to achieve three resolution levels $R_i = \{128, 512, 2048 | i = 1, 2, 3\}$, which correspond to the number of points of the attention output. At the same time, we generate a corresponding number of 2D grids and connect them together. Finally, the connected point cloud is converted into a 3D potential codewords through MLP. The dimensions corresponding to the point cloud of these three resolutions are $D_i = \{256, 128, 3 | i = 1, 2, 3\}$. The down-module aggregates these expanded points into detailed local features through concatenation and convolution operations. As shown in the pink box in Figure 1, the completion effect increases gradually through three folding operations, and the final output is a fine surface structure. The attention mechanism connects the encoder and the decoder. It looks for similar regions in the encoder and guides the decoder to reconstruct the missing regions with high quality.

3.4. Loss Function

Charmfer Distance (CD) and Earth Movement Distance (EMD) are used to define a loss function as the metric distance of point cloud completion.The total loss is the weighted sum of CD and EMD, defined as

$$\ell_{Total\_Loss} = \rho \ell_{CD\_Loss} + \ell_{EMD\_Loss}, \tag{1}$$

where ρ is a constant with ρ =10 [17]. The ℓ_{CD_Loss} can be calculated by

$$\ell_{CD\_Loss} = \frac{\sum_{x \in S_1,\ y \in S_2} \min \| x - y \|^2_2}{|S_1|} + \frac{\sum_{x \in S_2,\ y \in S_1} \min \| y - x \|^2_2}{|S_2|}, \tag{2}$$

where the first part of the ℓ_{CD_loss} is used to ensure the minimum distance between the completed point clouds and groundtruth (Gt) point clouds. The second part ensures the coverage of Gt point clouds on the completed point clouds. S_1 and S_2 represent two sets of point cloud models. x and y are the coordinates of sampling points in the point cloud models, respectively. The ℓ_{EMD_loss} can be calculated by

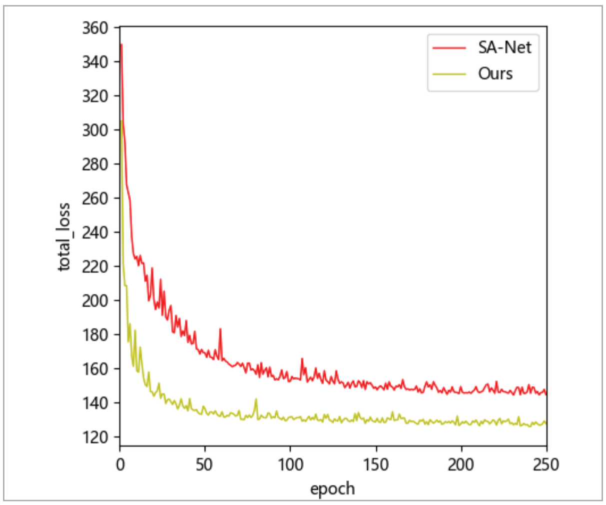
$$\ell_{EMD\_Loss}(S_1, S_2) = \min_{\phi: S_1 \to S_2} \sum_{x \in S_1} \| x - \phi(x) \|_2, \tag{3}$$

where ℓ_{EMD_Loss} is used to ensure the minimum distance between the completed point clouds and Gt point clouds. ϕ is the mapping function.

Regularization is a technique that promotes model complexity and accuracy. PointCutMix [36] uses reg-

Figure 4

Comparison of total loss during training between our proposed method and SA-Net



ularization to prevent the model from overfitting the training data by adding additional constraints or penalties during the training process. This improves the model's generalization ability, allowing it to perform better on new, unseen data. Figure 4 depicts the total loss training curve of our method and SA-Net using regularization. It can be seen that our method has a lower loss, which shows that our method has better completion effect.

4. Experimental Results

In this section, we will verify the effectiveness of our proposed network. We use the deep learning model built with the TensorFlow framework in Python. The operating system is Ubuntu 16.04 and the graphics card is NVIDIA Titan X GPU. The Adam optimizer is used to perform 250 training rounds for each type of point clouds. The training batch is 18. The initial learning rate is set to 0.0001. The number of neighbors of a sampling point is 12 and the model parameters with the best effect on the training set are saved for testing.

4.1. Dataset

ShapeNet dataset is a large-scale 3D shape dataset widely used in computer graphics, robotics, and computer vision. The ShapeNet dataset contains pairs of complete and partial point clouds, among which plane, cabinet, car, chair, lamp, couch, table, and boat are often used to test the performance of point cloud completion networks. The number of input and output points per object is 2048, but incomplete. Of course, the number of real points is also 2048.

4.2. Implementation Details

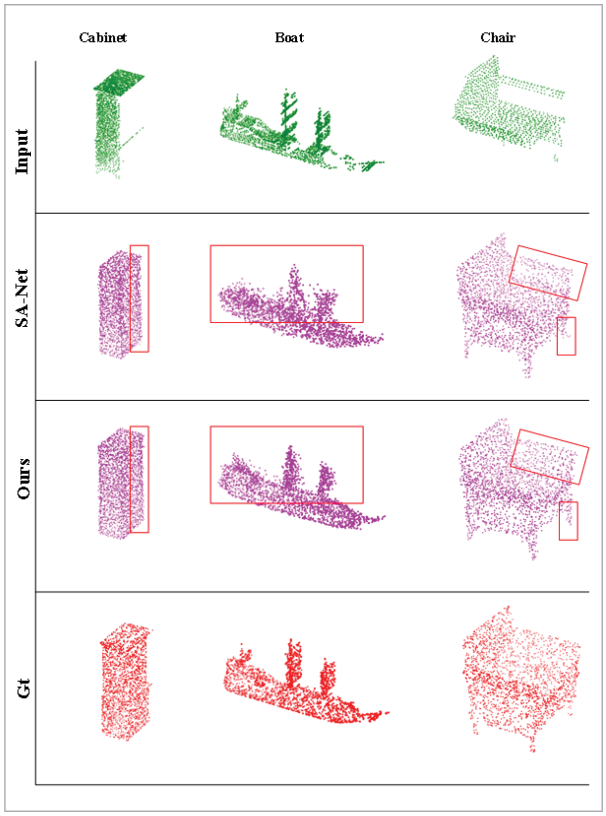
In Figure 5, we present the visualization results comparing our method with SA-Net on chair, cabinet and boat. It can be seen that the completion shape of our method is closer to the ground truth.

For example, when predicting the edges of the cabinet, sails and the arms and legs of the chair, our method can generate more realistic local structure, denser and more uniform point clouds, as well as reconstruct smoother surfaces and more precise geometry.

First, the presence of noise may have an impact on the accuracy of point cloud completion. Noise may make point cloud data more complex and confusing, causing the algorithm to be unable to accurately identify and restore the original point cloud structure during the completion process. Therefore, before performing point cloud completion, it is often necessary to denoise the point cloud data to reduce the impact of noise on the completion results. Secondly, the degree of damage to the point cloud will also affect the completion effect. Mildly damaged point cloud data may retain the original structure and characteristics to a large extent, so it is relatively easy for completion algorithms to restore the complete point cloud. However, if the point cloud data is highly damaged, it may make the original structure and features difficult to identify, thereby increasing the difficulty of completion. Under the above circumstances, it is very important for the point cloud completion network to have excellent noise suppression modules and feature extraction modules.

Figure 5

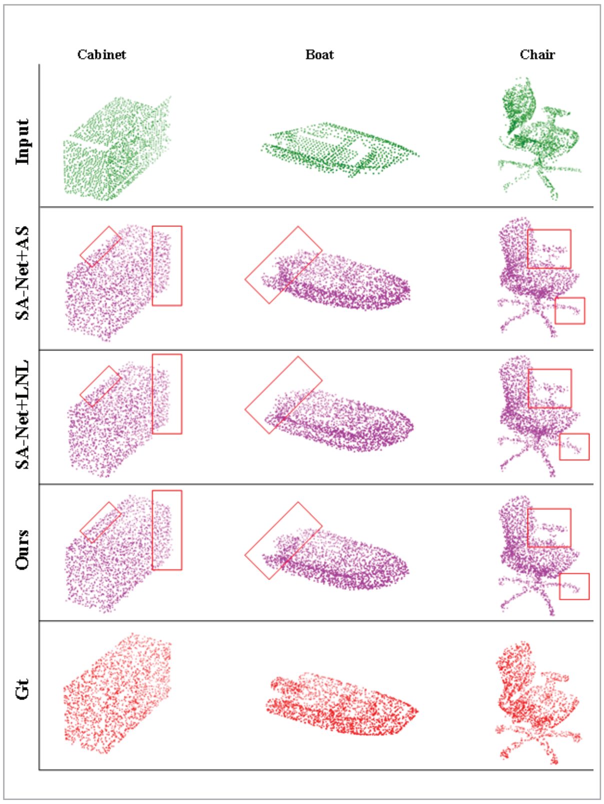
The visualization results comparing our method with SA-Net under Chair, Cabinet, and Boat. From top to bottom are input, SA-Net, our method and Gt



To further illustrate the effectiveness of the AS and LNL modules, we consider the following benchmark schemes. (1) SA-Net without the AS and LNL modules; (2) SA-Net+AS with only AS module; (3) SA-Net+LNL with only LNL module; (4) SA-Net+AS+LNL (Ours), including AS module and LNL module. From Figure 6, we can see that our method can get smooth cabinet edges, non-messy stern spots of boat, and chair legs, armrests. In short, a refined local geometric structure can be generated using both the AS module and the LNL module. In addition, Table 1 shows the comparison of parameter sizes of different methods. To generate denser, more uniform points, finer geometry, it needs to increase the number of parameters.

Figure 6

Comparison of ablation experiments



Next, we use the CD as an evaluation metric. Since the CD value is too small to be intuitive, we multiply it by 10^4 and reserve two decimal places as the test result. If the CD value is low, the distance between the probability distribution predicted by the model and the true label distribution is small, which means that the model has better classification performance on the training set. On the contrary, if the CD value is large, the classification performance of the model must poor.

Table 2 gives the ablation results. Compared with SA-Net, the CD of SA-Net+AS is reduced by 5.49% for cabinet (9.11 vs 8.61), 11.62% for boat (7.23 vs 6.39) and 1.23% for chair (8.94 vs 8.82), respectively. Furthermore, the CD of SA-Net+LNL is reduced by 4.39% for cabinet (9.11 vs 8.71), 7.47% for boat (7.23 vs 6.69) and 0.45% for chair (8.94 vs 8.90), respectively. These could prove the effectiveness of the AS module and LNL module.

In order to further verify the effectiveness of the proposed method, we compare with PCN [33], FoldingNet [31], AtlasNet [21], TopNet [23] and SA-Net

Table 1

The parameter sizes of different methods

Methods	SA-Net	SA-Net+AS	SA-Net+LNL	SA-Net+AS+LNL
Parameters(MB)	55.9	72.9	73.4	73.4

Table 2

The CD of different methods

Model	Cabinet	Boat	Chair
SA-Net	9.11	7.23	8.94
SA-Net+AS	8.61	6.39	8.82
SA-Net+LNL	8.71	6.69	8.90
SA-Net+AS+LNL	**8.56**	**6.30**	**8.79**

Table 3
Comparison of point cloud completion on ShapeNet dataset

Methods	Average	Plane	Cabinet	Car	Chair	Lamp	Couch	Table	Boat
AtlasNet	17.69	10.37	23.40	13.41	24.16	20.24	20.82	17.52	11.62
PCN	14.72	8.09	18.32	10.53	19.33	18.52	16.44	16.34	10.21
FoldingNet	16.48	11.18	20.15	13.25	21.48	19.19	19.09	17.80	10.69
TopNet	9.72	5.50	12.02	8.90	12.56	**9.54**	12.20	**9.57**	7.51
SA-Net	7.74	**2.18**	9.11	**5.56**	8.94	9.98	7.83	9.94	7.23
Ours	**7.45**	2.29	**8.56**	5.70	**8.79**	10.01	**7.83**	10.14	**6.30**

[25]. The comparison results are shown in Table 3. Compared with SA-Net, our method achieves 6.04% for cabinet (8.56 vs 9.11), 13.29% for boat (6.30 vs 7.23) and 1.71% for chair (8.79 vs 8.94), respectively. It achieves an average CD reduction of 3.50% (7.45 vs 7.74). We can see that our method outperforms others in many categories. Note that the results of all methods are referenced from the public leaderboards of the ShapeNet benchmark.

5. Conclusion

In this paper, the characteristics of our method are based on the ability to suppress noise and enhance point cloud feature extraction, thereby improving the robustness and accuracy of point cloud processing. Thereforer, we propose an adaptive sampling non-local neural network to complete the point clouds. The effect of outliers is mitigated by the PointASNL encoder, which adjusts the information of sampling points in real time by weighting adjacent points. Different from the traditional methods, we also consider the correlation between the sampling points and the entire points in a certain level. In this way, the feature information of points cloud can be extracted and aggregated more accurately. Experiments show the completion effects of multiple categories and prove the effectiveness of AS and LNL modules. Quantitative and qualitative results on the ShapeNet data set show that compared with SA-Net, our method has a significant reduction in the chamfer distance and average chamfer distance of multiple categories and the reconstructed point cloud model has more detailed local areas. Our method uses max pooling to aggregate features, which leads to information loss during feature extraction and requires further improvement. The point cloud completion effect in complex scenes may not be ideal and has certain limitations.

Acknowledgments

This work was supported in part by the R&D Program of Beijing Municipal Education Commission (KM202310009001), in part by the Scientific Research Foundation of North China University of Technology (110051360002), and in part by the National Key Research and Development Program of China (2020YFC0811004).

References

1. Chang, Y., Jung, C., Xu, Y. FinerPCN: High fidelity Point Cloud Completion Network Using Pointwise Convolution. Neurocomputing, 2021, 460, 266-276. https://doi.org/10.1016/j.neucom.2021.06.080
2. Chen, X., Chen, B., Mitra, N. J. Unpaired Point Cloud Completion on Real Scans Using Adversarial Training. ArXiv preprint ArXiv:1904.00069, 2019.
3. Creswell, A., White, T., Dumoulin, V. Arulkumaran, K., Sengupta, B., Bharath, A. A. Generative Adversarial Networks: An Overview.IEEE Signal Processing Magazine 2018, 35, 53-65. https://doi.org/10.1109/MSP.2017.2765202
4. Feng, Y., You, H., Zhang, Z., Ji, R., Gao, Y. Hypergraph Neural Networks. Proceedings of the AAAI Conference

on Artificial Intelligence, 2019, 33, 3558-3565. https://doi.org/10.1609/aaai.v33i01.33013558

5. Gao, J., Lan, J., Wang, B., Li, F. SDANet: Spatial Deep Attention-based for Point Cloud Classification and Segmentation. Machine Learning, 2022, 111, 1327-1348. https://doi.org/10.1007/s10994-022-06148-1
6. Groueix, T., Fisher, M., Kim, V. G., Aubry, M. A Papier-mâché Approach to Learning 3d Surface Generation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 216-224. https://doi.org/10.1109/CVPR.2018.00030
7. Guo, M. H., Cai, J. X., Liu, Z. N., Mu, T. J., Martin, R. R., Hu, S. M. Pct: Point Cloud Transformer. Computational Visual Media, 2021, 7, 187-199. https://doi.org/10.1007/s41095-021-0229-5
8. He, X., Zhou, Y., Zhou, Z., Bai, S., Bai, X. Triplet-center Loss for Multi-view 3d Object Retrieval. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 1945-1954. https://doi.org/10.1109/CVPR.2018.00208
9. Huang, Z., Yu, Y., Xu, J., Ni, F., Le, X. Pf-net: Point Fractal Network for 3d Point Cloud Completion. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, 7662-7670. https://doi.org/10.1109/CVPR42600.2020.00768
10. Jing, W., Zhang, W., Li, L., Di, D., Chen, G.,Wang, J.AG-Net: An Attention-based Graph Network for Point Cloud Classification and Segmentation. Remote Sensing, 2022, 14, 1036. https://doi.org/10.3390/rs14041036
11. Kanezaki, A., Matsushita, Y., Nishida, Y. Rotationnet: Joint Object Categorization and Pose Estimation Using Multiviews from Unsupervised Viewpoints. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 5010-5019. https://doi.org/10.1109/CVPR.2018.00526
12. Kulikajevas, A., Maskeliūnas, R., Damaševičius, R., Misra, S. Reconstruction of 3D Object Shape Using Hybrid Modular Neural Network Architecture Trained on 3D Models from ShapeNetCore Dataset. Sensors, 2019, 19, 1553. https://doi.org/10.3390/s19071553
13. Li, Y., Bu, R., Sun, M., Wu, W., Di, X., Chen, B. Pointcnn: Convolution on X-transformed Points. Advances in Neural Information Processing Systems, 2018, 31.
14. Litany, O., Bronstein, A., Bronstein M., Makadia, A. Deformable Shape Completion with Graph Convolutional Autoencoders. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 1886-1895. https://doi.org/10.1109/CVPR.2018.00202
15. Liu, Y., Fan, B., Xiang, S., Pan, C. Relation-shape Convolutional Neural Network for Point Cloud Analysis. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, 8895-8904. https://doi.org/10.1109/CVPR.2019.00910
16. Liu, M., Sheng, L., Yang, S., Shao, J., Hu, S. M. Morphing and Sampling Network for Dense Point Cloud Completion. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 11596-11603. https://doi.org/10.1609/aaai.v34i07.6827
17. Pan, X., Xia, Z., Song, S., Li, L. E., Huang, G. 3d Object Detection with Pointformer. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, 7463-7472. https://doi.org/10.1109/CVPR46437.2021.00738
18. Pierdicca, R., Paolanti, M., Matrone, F., Martini, M., Morbidoni, C., Malinverni, E. S., Frontini, E., Lingua, A. M. Point Cloud Semantic Segmentation Using a Deep Learning Framework for Cultural Heritage. Remote Sensing, 2020, 12, 1005. https://doi.org/10.3390/rs12061005
19. Qi, C. R., Su, H., Mo, K., Guibas, L. J. Pointnet: Deep Learning on Point Sets for 3d Classification and Segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, 652-660. https://doi.org/10.1109/CVPR.2017.16
20. Qi, C. R., Yi, L., Su, H., Guibas, L. J. Pointnet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. Advances in Neural Information Processing Systems, 2017, 30.
21. Ryselis, K., Blažauskas, T., Damaševičius, R., Maskeliūnas, R. Computer-aided Depth Video Stream Masking Framework for Human Body Segmentation in Depth Sensor Images. Sensors, 2022, 22, 3531. https://doi.org/10.3390/s22093531
22. Sarmad, M., Lee, H. J., Kim, Y. M. Rl-gan-net: A Reinforcement Learning Agent Controlled Gan Network for Real-time Point Cloud Shape Completion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, 5898-5907 https://doi.org/10.1109/CVPR.2019.00605
23. Tchapmi, L. P., Kosaraju, V., Rezatofighi, H., Reid, I., Savarese, S. Topnet: Structural Point Cloud Decoder. Proceedings of the IEEE/CVF Coference on Computer Vision and Pattern Recognition, 2019, 383-392. https://doi.org/10.1109/CVPR.2019.00047
24. Wang, N., Zhang, Y., Li, Z., Fu, Y., Liu, W., Jiang, Y. G. Pixel2mesh: Generating 3d Mesh Models From Single Rgb Images. Proceedings of the European Conference

on Computer Vision (ECCV), 2018, 52-67. https://doi.org/10.1007/978-3-030-01252-6_4

25. Wen, X., Li, T., Han, Z., Liu, Y. S. Point Cloud Completion by Skip-attention Network with Hierarchical Folding. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, 1939-1948. https://doi.org/10.1109/CVPR42600.2020.00201
26. Wen, X., Han, Z., Cao, Y. P., Wan, P., Zheng, W., Liu, Y. S. Cycle4completion: Unpaired Point Cloud Completion Using Cycle Transformation with Missing Region Coding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, 13080-13089. https://doi.org/10.1109/CVPR46437.2021.01288
27. Wen, X., Xiang, P., Han, Z., Cao, Y. P., Wan, P., Zheng, W., Liu, Y. S. PMP-Net++: Point cloud Completion by Transformer-enhanced Multi-step Point Moving Paths. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45, 852-867. https://doi.org/10.1109/TPAMI.2022.3159003
28. Wen, X., Xiang, P., Han, Z., Cao, Y. P., Wan, P., Zheng, W., Liu, Y. S. Pmp-net: Point Cloud Completion by Learning Multi-step Point Moving Paths. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, 7443-7452. https://doi.org/10.1109/CVPR46437.2021.00736
29. Xia, Y., Xu, Y., Wang, C., Stilla, U. VPC-Net: Completion of 3D Vehicles from MLS Point Clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 174, 166-181. https://doi.org/10.1016/j.isprsjprs.2021.01.027
30. Yan, X., Zheng, C., Li, Z., Wang, S., Cui, S. Pointasnl: Robust Point Clouds Processing Using Nonlocal Neural Networks with Adaptive Sampling. Proceedings of the IEE Networks with Adaptive Sampling. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020, 5589-5598. https://doi.org/10.1109/CVPR42600.2020.00563
31. Yang, Y., Feng, C., Shen, Y., Tian, D. Foldingnet: Point Cloud Auto-encoder Via Deep Grid Deformation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018, 206-215. https://doi.org/10.1109/CVPR.2018.00029
32. Yu, T., Meng, J., Yuan, J. Multi-view Harmonized Bilinear Network for 3d Object Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 186-194. https://doi.org/10.1109/CVPR.2018.00027
33. Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M. Pcn: Point Completion Network. In 2018 International Conference on 3D Vision (3DV), 2018, 728-737. https://doi.org/10.1109/3DV.2018.00088
34. Zhang, J., Chen, X., Cai, Z., Pan, L., Zhao, H., Yi, S., Yeo, C. K., Dai, B., Loy, C. C. Unsupervised 3d Shape Completion Through Gan Inversion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, 1768-1777. https://doi.org/10.1109/CVPR46437.2021.00181
35. Zhang, W., Yan, Q., Xiao, C. Detail Preserved Point Cloud Completion Via Separated Feature Aggregation. Proceedings of the European Conference on Computer vision (ECCV), 2020, 512-528. https://doi.org/10.1007/978-3-030-58595-2_31
36. Zhang, J., Chen, L., Ouyang, B., Liu, B., Zhu, J., Chen, Y., Meng, Y., Wu, D. Pointcutmix: Regularization Strategy for Point Cloud Classification.Neurocomputing, 2022, 505, 58-67. https://doi.org/10.1016/j.neucom.2022.07.049
37. Zhang, X., Fu, C., Zhao, Y., Xu, X. Hybrid Feature CNN Model for Point Cloud Classification and Segmentation. IET Image Processing, 2020, 14, 4086-4091. https://doi.org/10.1049/iet-ipr.2020.0658
38. Zhao, H., Jiang, L., Jia, J., Torr, P. H., Koltun, V. Point Transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, 16259-16268. https://doi.org/10.1109/ICCV48922.2021.01595
39. Zhou, Y., Tuzel, O. Voxelnet: End-to-end Learning for Point Cloud Based 3d Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 4490-4499. https://doi.org/10.1109/CVPR.2018.00472