PointMLP

: Rethinking Network Design and Local Geometry in Point Cloud

A SIMPLE MLP FRAMEWORK

Seungyeol Lee

CHAPTER

Part 1 Purpose of Research

- Point Cloud Classification & Part Segmentation

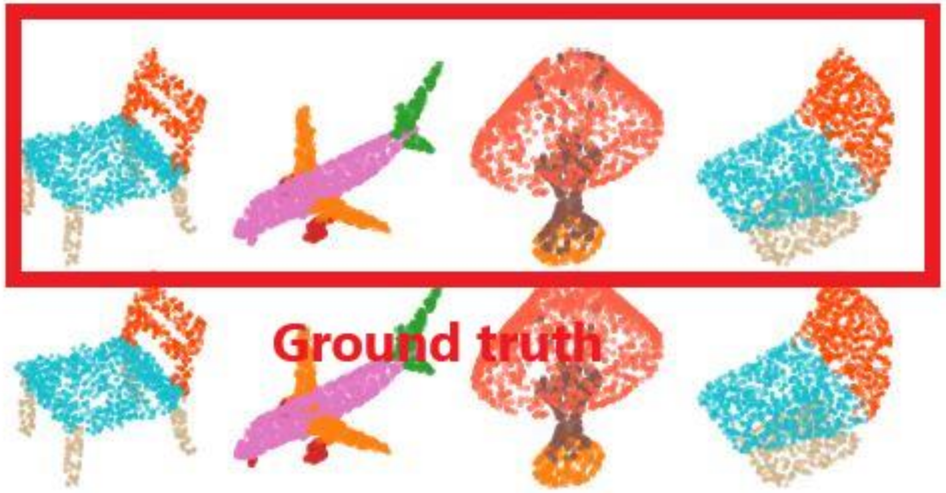


Figure 5: Part segmentation results on ShapeNetPart. Top line is ground truth and bottom line is our prediction.

Part 1
Purpose of Research

- Point Cloud Classification & Part Segmentation

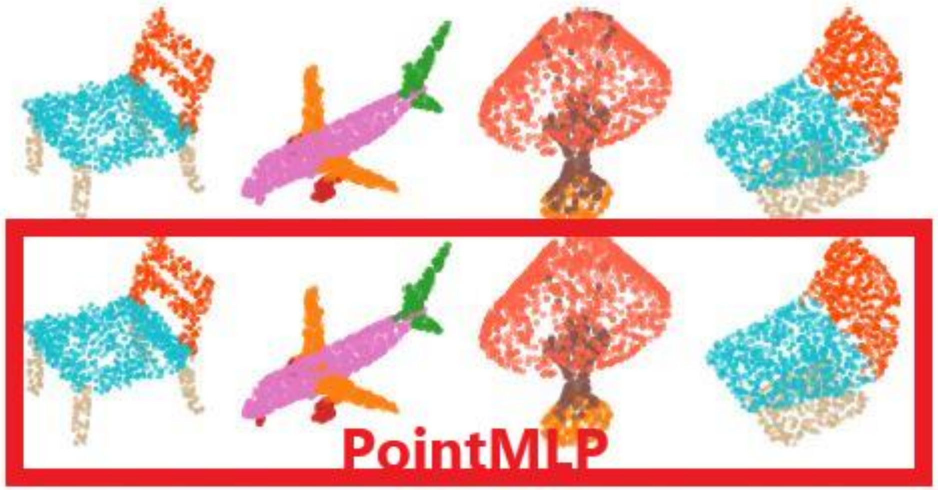


Figure 5: Part segmentation results on ShapeNetPart. Top line is ground truth and bottom line is our prediction.

Part 1

Purpose of Research

- Point Cloud Classification & Part Segmentation

Figure 5: Part segmentation results on ShapeNetPart. Top line is ground truth and bottom line is our prediction.

Part 2 **Limitations of Previous Research**

1. Point Cloud Analysis

- VoxNet : project the original point clouds to intermediate voxels

- PointNet : process the original point cloud sets directly

- PointNet++ : introduce a hierarchical feature learning paradigm

Part 2

Limitations of Previous Research

2. Local Geometry Exploration

Development of Local Geometry Exploration

-> Minimal Improvements

Part 2 **Limitations of Previous Research**

2. Local Geometry Exploration

(1) Convolution

PointConv : approximate continuous weight and density functions in convolutional filters using an MLP

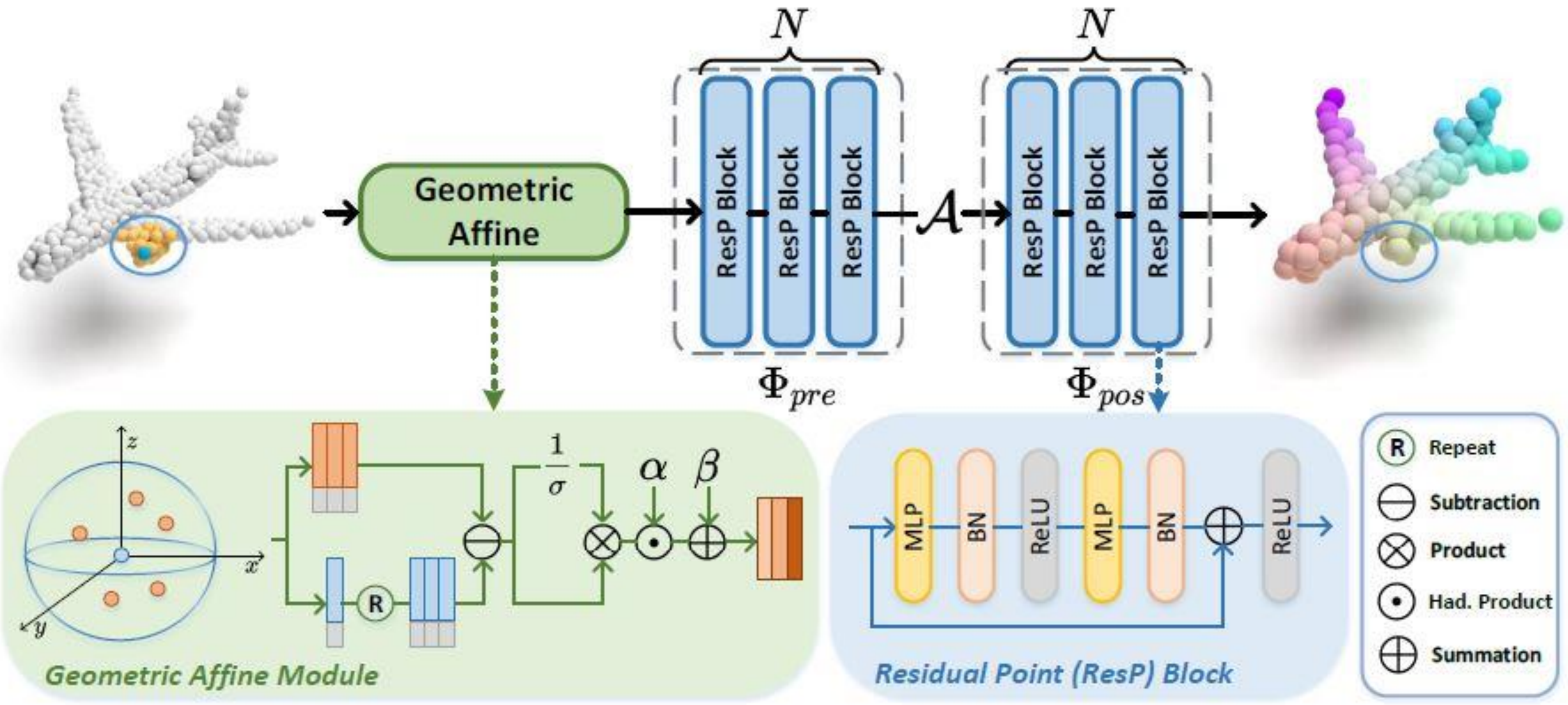
(2) Graph

EdgeConv : generate edge features that describe the relationships between a point and its neighbors

(3) Attention

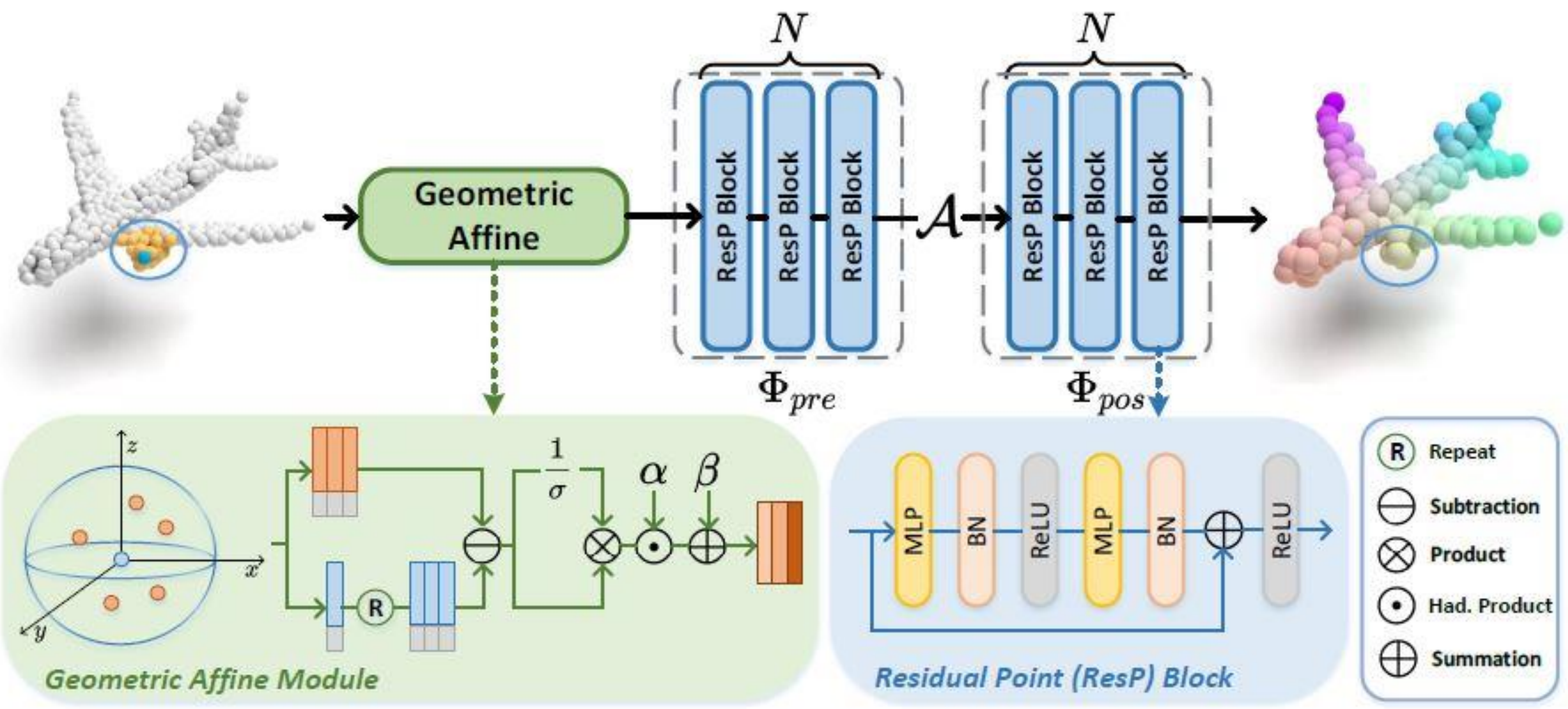
Point Transformer : exhibit excellent ability on relationship exploration

Part 3 **PointMLP**



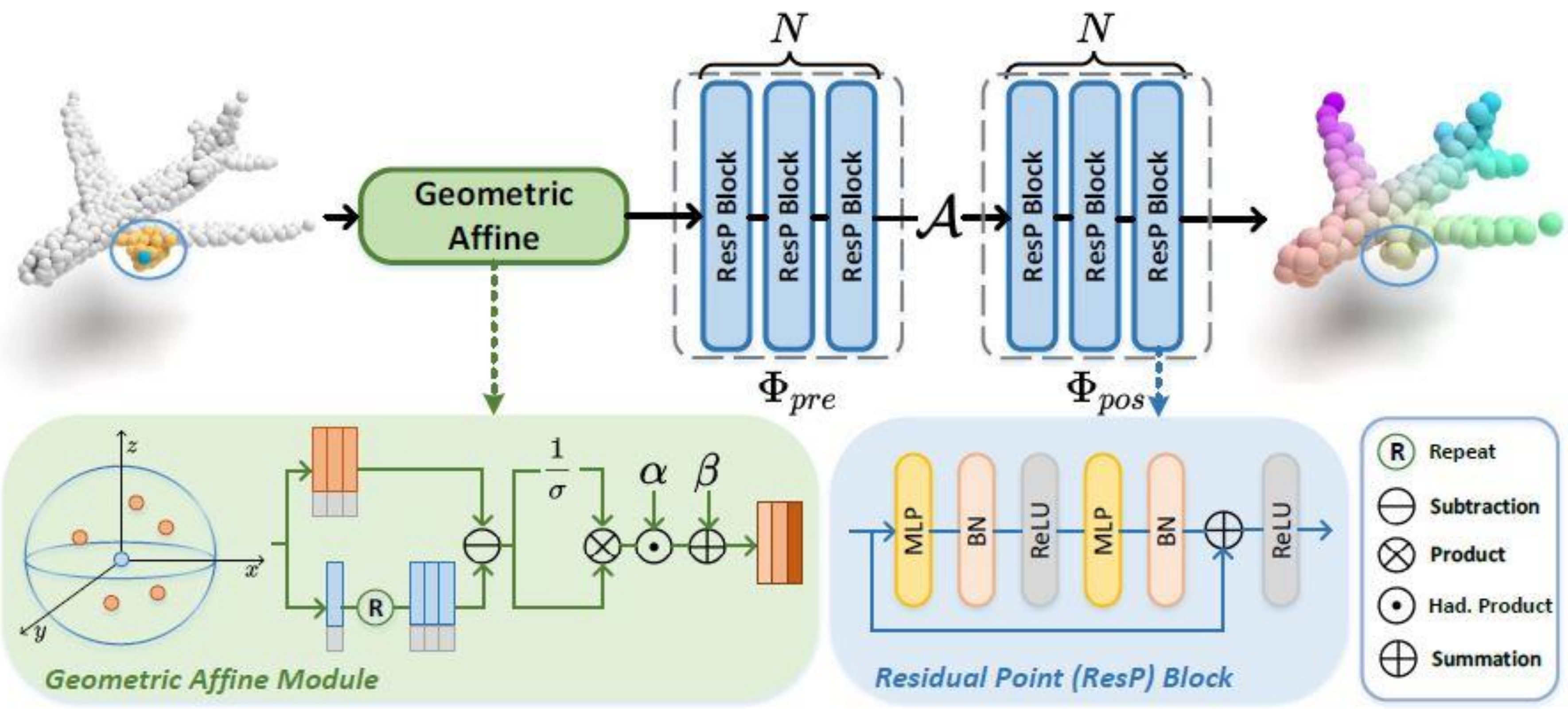
Rethinking the necessity of "sophisticated"local feature extractors

Part 3
PointMLP



Revisiting the succinct design philosophy in point cloud analysis

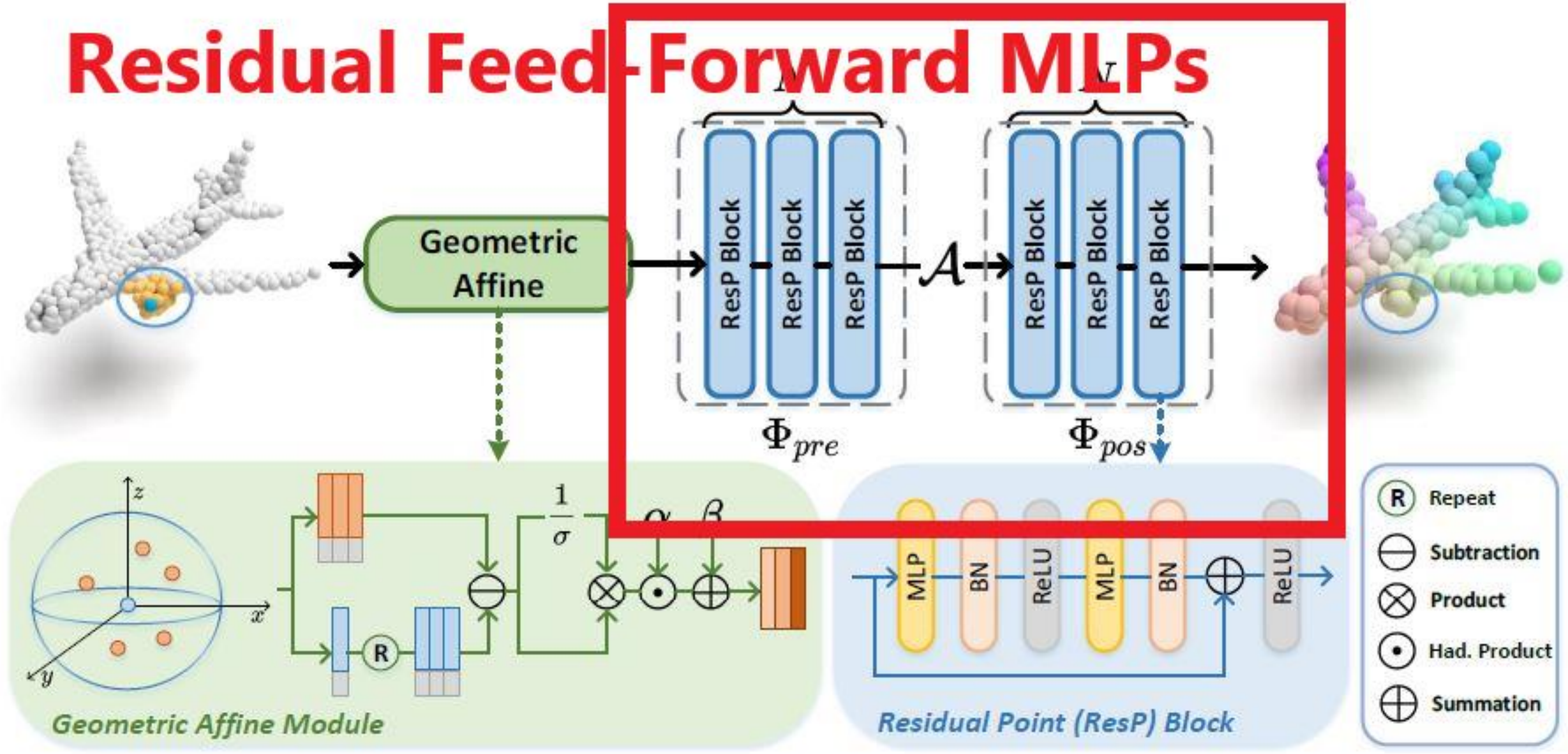
Part 3 # PointMLP



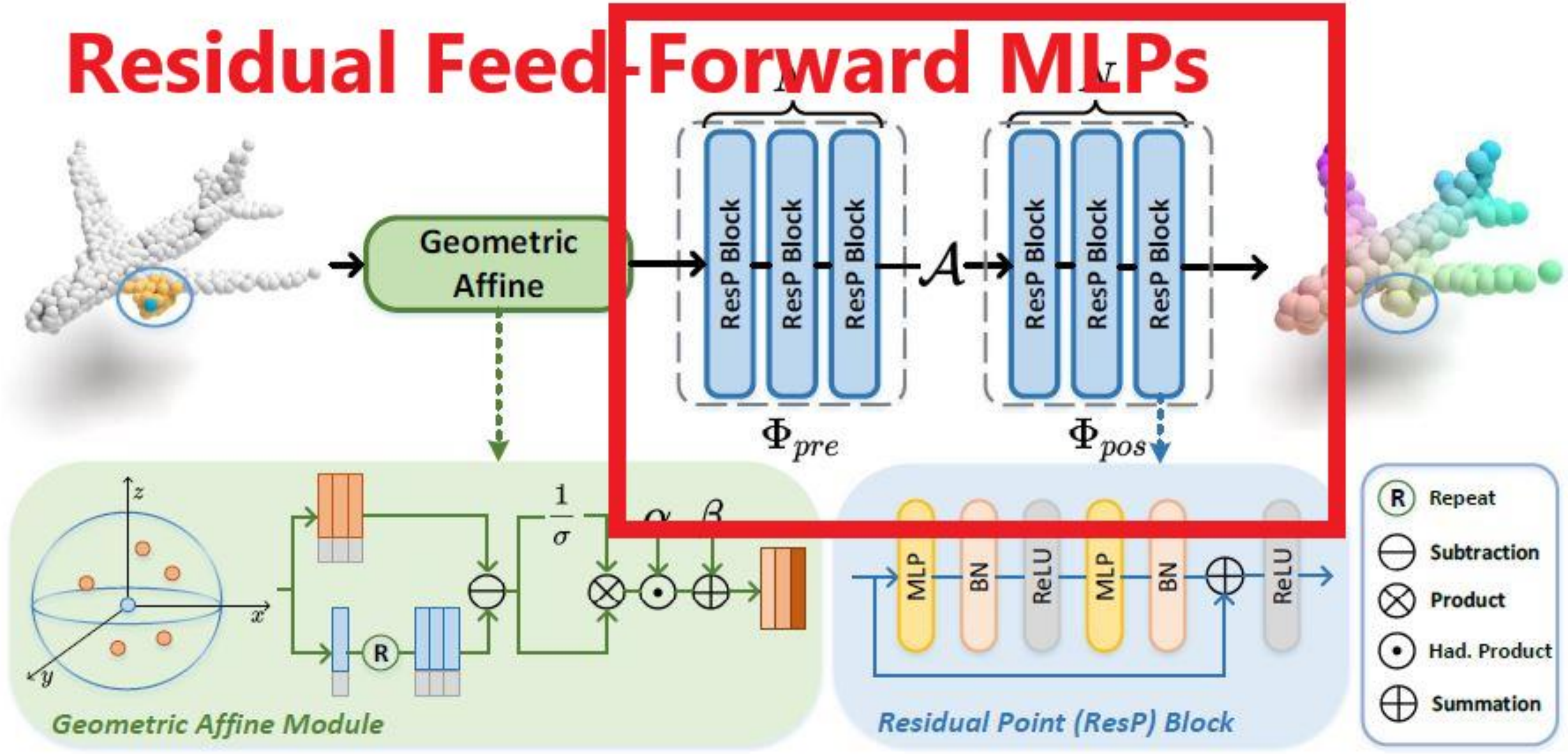
Pure Residual MLP Network

Part 4

Key Components of PointMLP



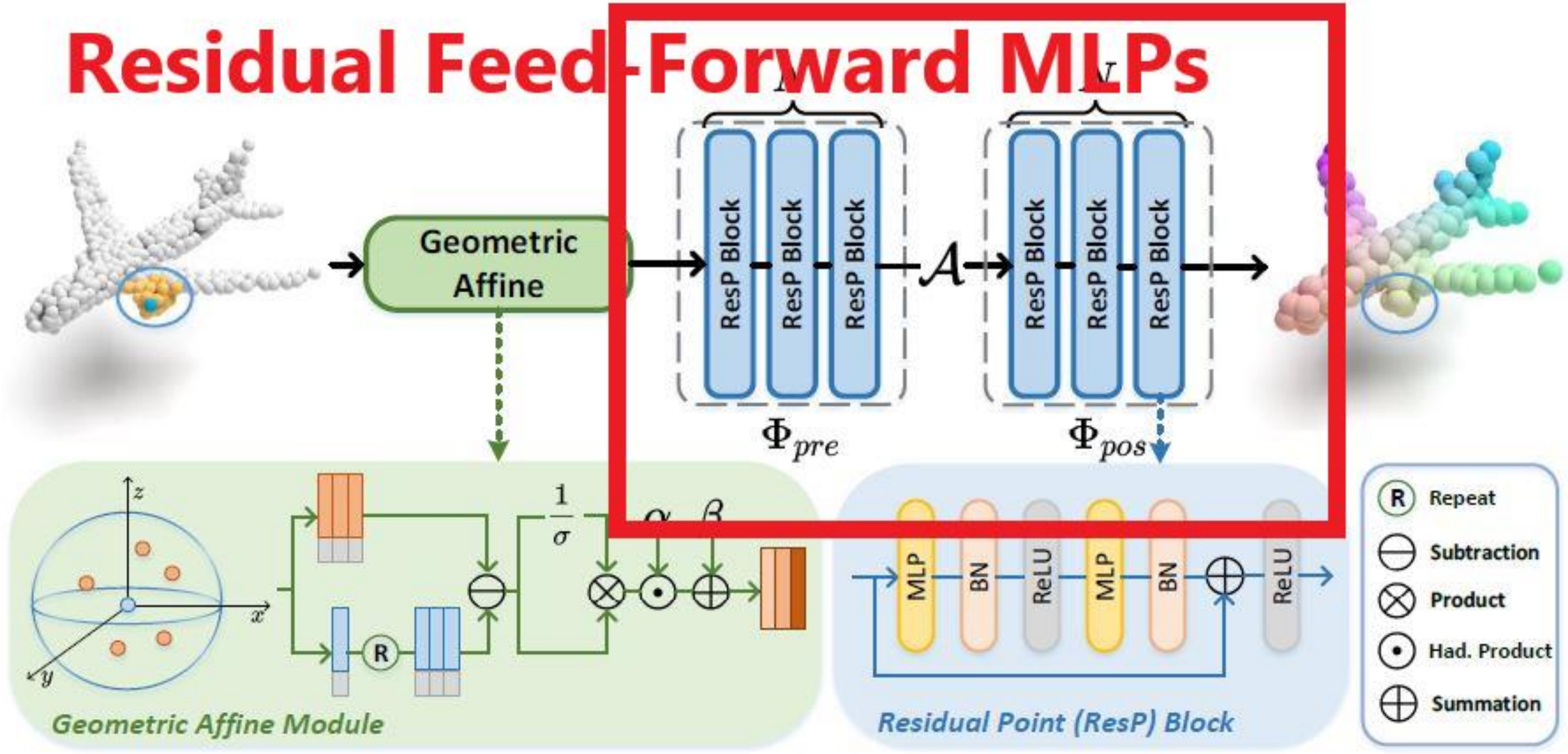
Part 4 Key Components of PointMLP



Φ_{pre} - **Residual MLP Blocks designed to learn shared weights from a local region**

Part 4

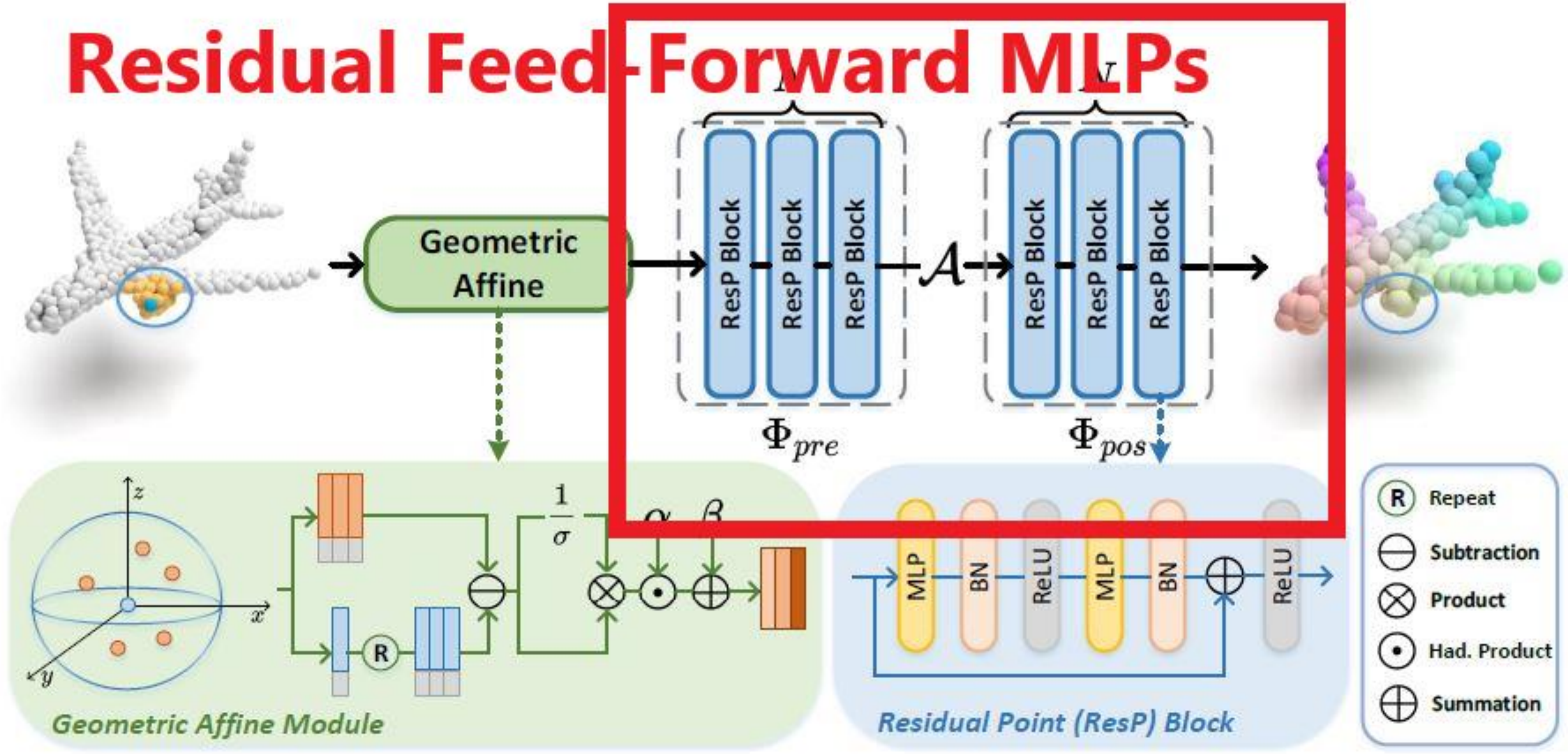
Key Components of PointMLP



Φ_{pos} - **Residual MLP Blocks designed to leverage extract deep aggregated features**

Part 4

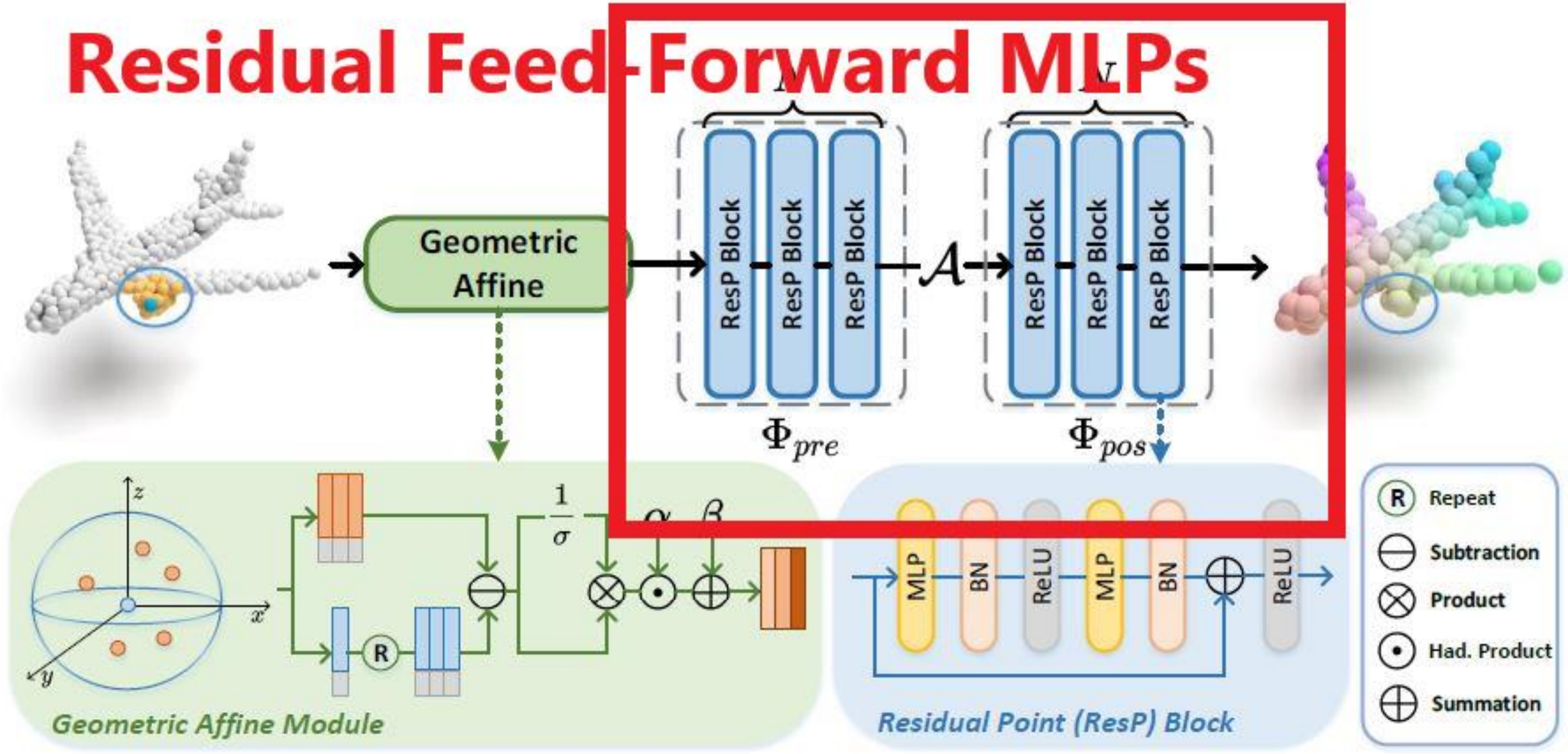
Key Components of PointMLP



$\mathcal{A}$ - Aggregation Function : max-pooling operation (like PointNet++)

Part 4

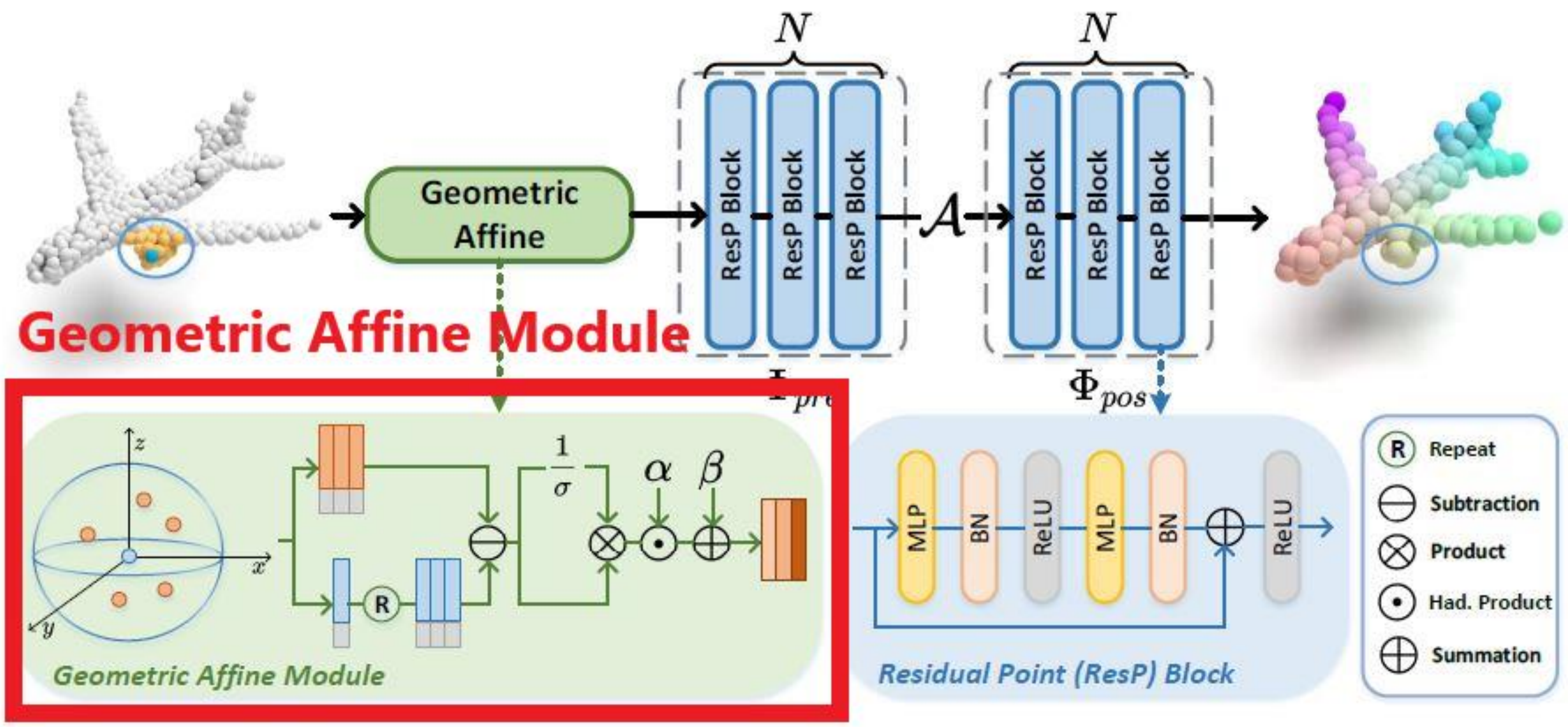
Key Components of PointMLP



$$g_i = \Phi_{pos}\left(\mathcal{A}\left(\Phi_{pre}\left(f_{i,j}\right), |j = 1, \cdots, K\right)\right)$$

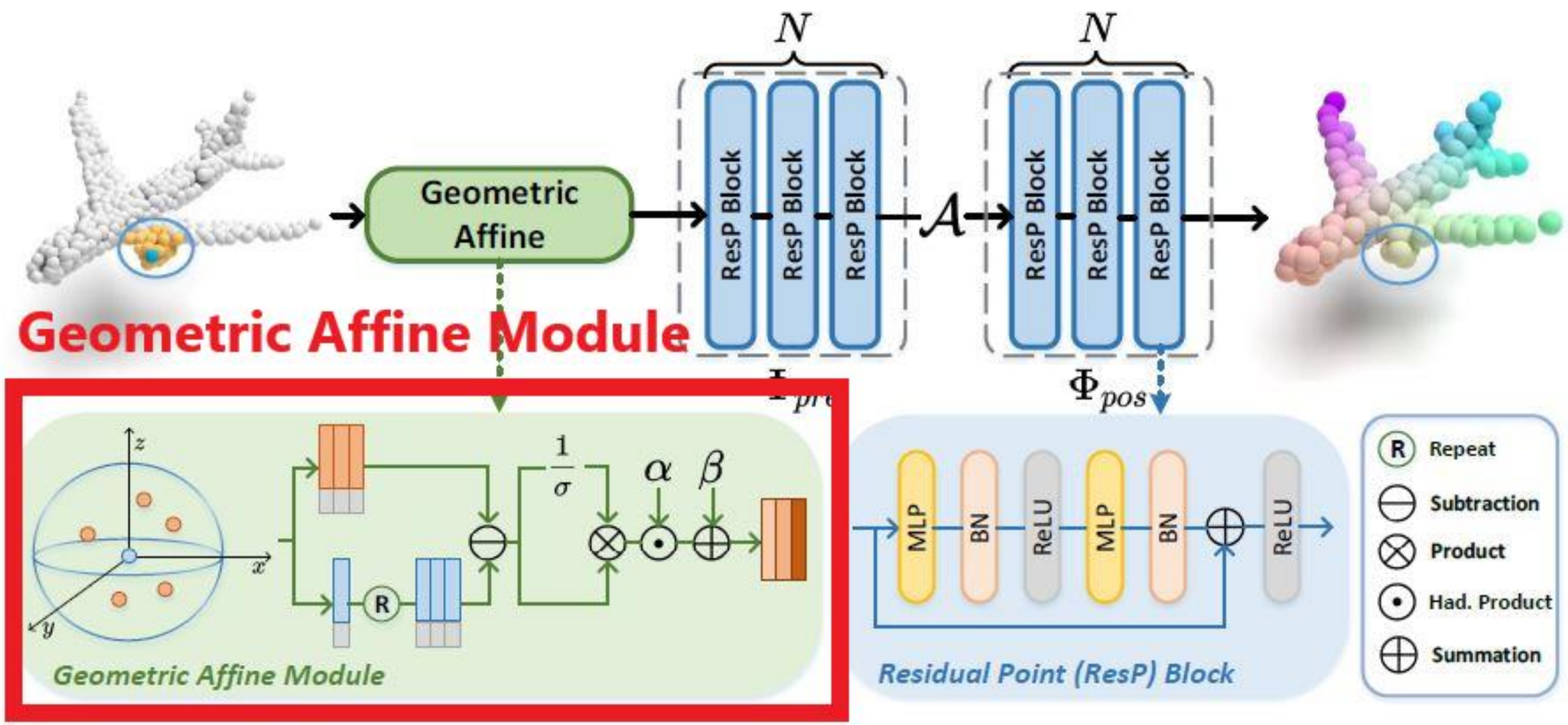
Part 4

Key Components of PointMLP



Part 4

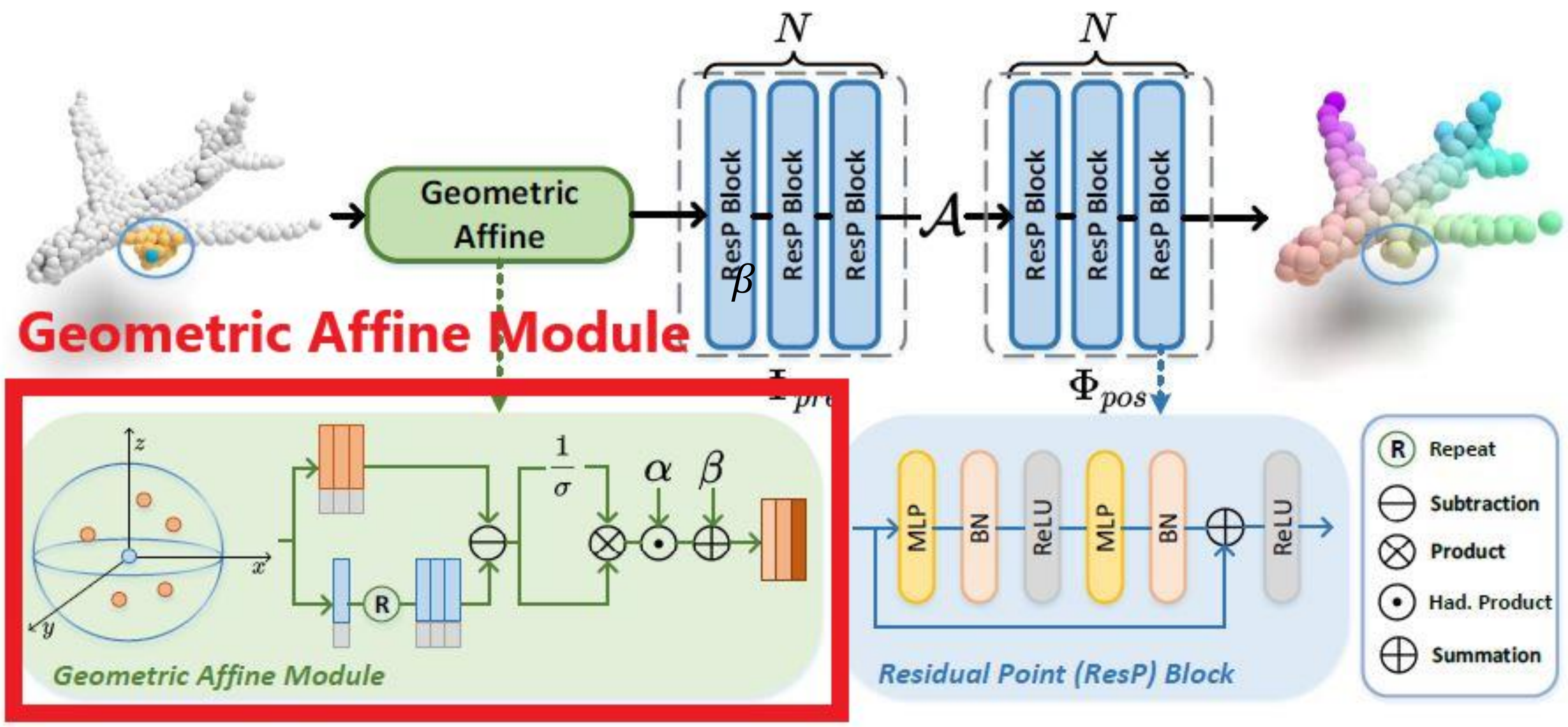
Key Components of PointMLP



$f_{i,j}$ **- Each neighbor point that is a $d-$dimensional vector**

$\{f_{i,j}\}_j$ **- Grouped local neighbors f_i of containing k points $(j = 1, 2, \dots, k)$**

Key Components of PointMLP

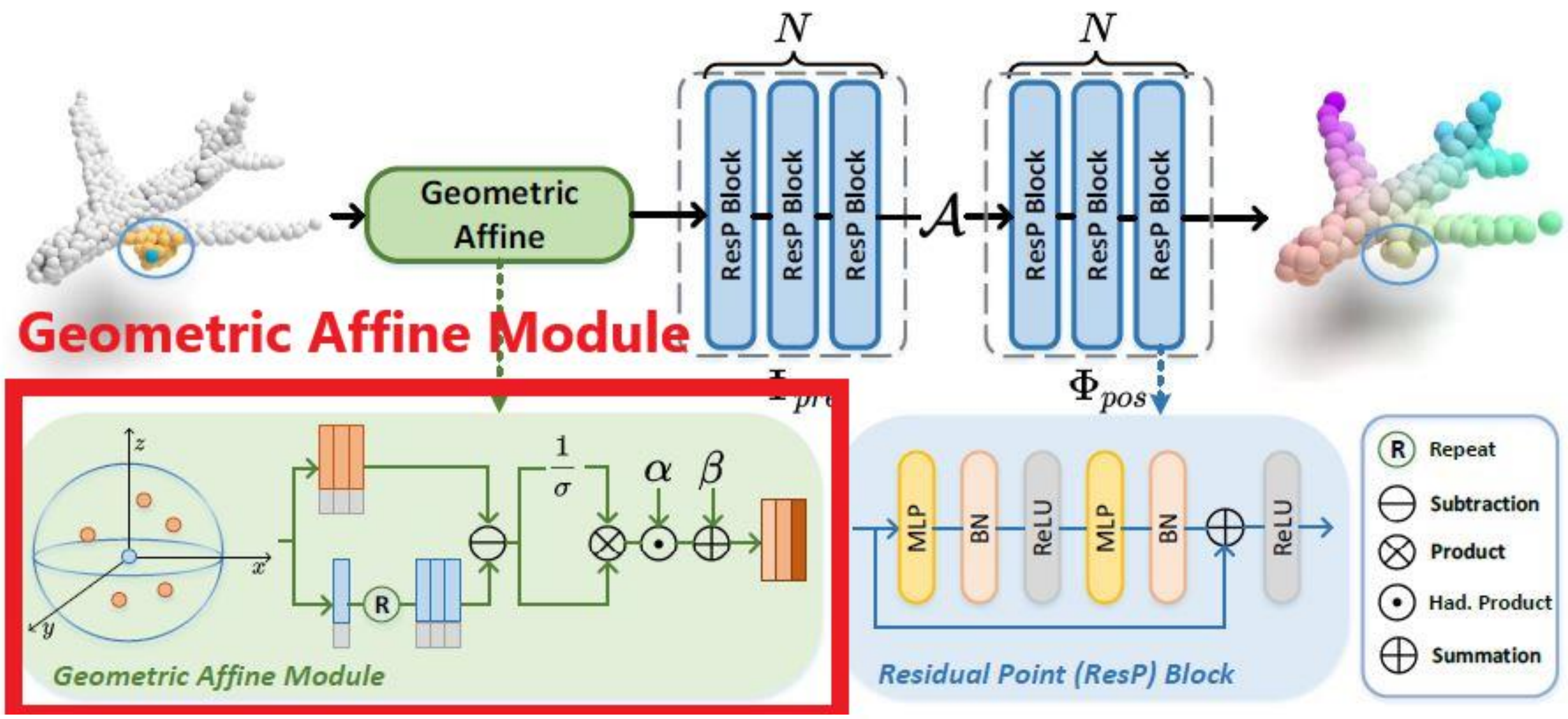


α, β - **Learnable parameters**

σ - **A scalar describing the feature derivation across all groups and channels**

Part 4

Key Components of PointMLP



$$\{f_{i,j}\} = \alpha \odot \frac{\{f_{i,j}\} - f_i}{\sigma + \epsilon} + \beta, \qquad \sigma = \sqrt{\frac{1}{k \times n \times d} \sum_{i=1}^{n} \sum_{j=1}^{k} (f_{i,j} - f_i)^2},$$

Part 4 PointMLP-elite

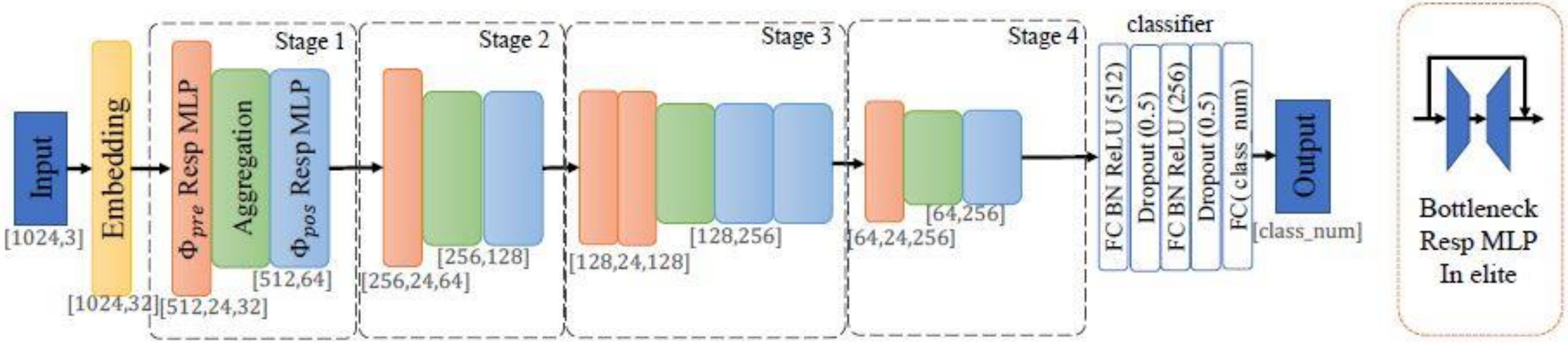


Figure 7: Detail architecture of PointMLP-elite for classification.

Less than 0.7M parameters and prominent inference speed

Experiments

ModelNet40

(9,843 training / 2,468 testing meshed CAD models / 40 categories)

Experiments

ScanObjectNN

(15,000 objects / 15 categories with 2,902 unique object instances)

Table 3: Classification results on ScanObjectNN dataset. We examine all methods on the most challenging variant (PB_T50_RS). For our pointMLP and PointMLP-elite, we train and test for four runs and report mean ± std results.

Method	mAcc(%)	OA(%)
3DmFV	58.1	63
PointNet (Qi et al., 2017a)	63.4	68.2
SpiderCNN (Xu et al., 2018)	69.8	73.7
PointNet++ (Qi et al., 2017b)	75.4	77.9
DGCNN (Wang et al., 2019)	73.6	78.1
PointCNN (Li et al., 2018b)	75.1	78.5
BGA-DGCNN (Uy et al., 2019)	75.7	79.7
BGA-PN++ (Uy et al., 2019)	77.5	80.2
DRNet (Qiu et al., 2021a)	78.0	80.3
GBNet (Qiu et al., 2021b)	77.8	80.5
SimpleView (Goyal et al., 2021)	-	80.5±0.3
PRANet (Cheng et al., 2021)	79.1	82.1
MVTN (Hamdi et al., 2021)	-	82.8
PointMLP (ours)	**83.9±0.5**	**85.4±0.3**
PointMLP-elite (ours)	**81.8±0.8**	**83.8±0.6**

Part 5

Experiments

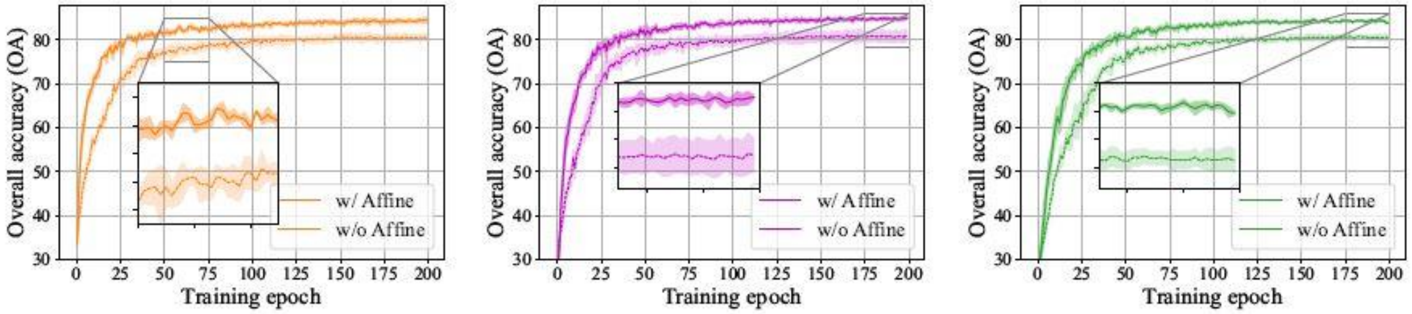
ShapeNetPart

(16,881 shapes / 16 classes belonging to 50 parts labels)

Table 6: Part segmentation results on the ShapeNetPart dataset. Empirically, our method is much faster than the best method KPConv, and presents a competitive performance.

Method	Cls. mIoU	Inst. mIoU	aero	bag	cap	car	chair	aerp-hone	guitar	knife	lamp	laptop	motor-bike	mug	pistol	rocket	skate-board	table
PointNet	80.4	83.7	83.4	78.7	82.5	74.9	89.6	73.0	91.5	85.9	80.8	95.3	65.2	93.0	81.2	57.9	72.8	80.6
PointNet++	81.9	85.1	82.4	79.0	87.7	77.3	90.8	71.8	91.0	85.9	83.7	95.3	71.6	94.1	81.3	58.7	76.4	82.6
Kd-Net	-	82.3	80.1	74.6	74.3	70.3	88.6	73.5	90.2	87.2	81.0	94.9	57.4	86.7	78.1	51.8	69.9	80.3
SO-Net	-	84.9	82.8	77.8	88.0	77.3	90.6	73.5	90.7	83.9	82.8	94.8	69.1	94.2	80.9	53.1	72.9	83.0
PCNN	81.8	85.1	82.4	80.1	85.5	79.5	90.8	73.2	91.3	86.0	85.0	95.7	73.2	94.8	83.3	51.0	75.0	81.8
DGCNN	82.3	85.2	84.0	83.4	86.7	77.8	90.6	74.7	91.2	87.5	82.8	95.7	66.3	94.9	81.1	63.5	74.5	82.6
P2Sequence	-	85.2	82.6	81.8	87.5	77.3	90.8	77.1	91.1	86.9	83.9	95.7	70.8	94.6	79.3	58.1	75.2	82.8
PointCNN	84.6	86.1	84.1	86.5	86.0	80.8	90.6	79.7	92.3	88.4	85.3	96.1	77.2	95.2	84.2	64.2	80.0	83.0
PointASNL	-	86.1	84.1	84.7	87.9	79.7	92.2	73.7	91.0	87.2	84.2	95.8	74.4	95.2	81.0	63.0	76.3	83.2
RS-CNN	84.0	86.2	83.5	84.8	88.8	79.6	91.2	81.1	91.6	88.4	86.0	96.0	73.7	94.1	83.4	60.5	77.7	83.6
SynSpec	82.0	84.7	81.6	81.7	81.9	75.2	90.2	74.9	93.0	86.1	84.7	95.6	66.7	92.7	81.6	60.6	82.9	82.1
SPLATNet	83.7	85.4	83.2	84.3	89.1	80.3	90.7	75.5	92.1	87.1	83.9	96.3	75.6	95.8	83.8	64.0	75.5	81.8
SpiderCNN	82.4	85.3	83.5	81.0	87.2	77.5	90.7	76.8	91.1	87.3	83.3	95.8	70.2	93.5	82.7	59.7	75.8	82.8
KPConv	85.1	86.4	84.6	86.3	87.2	81.1	91.1	77.8	92.6	88.4	82.7	96.2	78.1	95.8	85.4	69.0	82.0	83.6
PA-DGC	84.6	86.1	84.3	85.0	90.4	79.7	90.6	80.8	92.0	88.7	82.2	95.9	73.9	94.7	84.7	65.9	81.4	84.0
PointMLP	84.6	86.1	83.5	83.4	87.5	80.54	90.3	78.2	92.2	88.1	82.6	96.2	77.5	95.8	85.4	64.6	83.3	84.3

Ablation Studies



(a) PointMLP 24-Layers (b) PointMLP 40-Layers (c) PointMLP 56-Layers

Figure 3: Four run results (mean ± std) of PointMLP with/without our geometric affine module on ScanObjectNN test set. We zoom in on the details of PointMLP40 to show the stability difference.

Table 4: Classification accuracy of pointMLP on ScanObjectNN test set using 24, 40, and 56 layers, respectively.

Depth	mAcc(%)	OA(%)
24 layers	83.4±0.4	84.8±0.5
40 layers	**83.9±0.5**	**85.4±0.3**
56 layers	83.2±0.2	85.0±0.1

Table 5: Component ablation studies on ScanObjectNN test set.

Φ_{pre}	Φ_{pos}	Affine	mAcc(%)	OA(%)
✗	✓	✓	80.8±0.4	82.8±0.0
✓	✗	✓	83.3±0.3	84.7±0.2
✓	✓	✗	79.1±1.7	81.5±1.4
✓	✓	✓	**83.9**±0.5	**85.4±0.3**

Part 6

Conclusions

- Simple yet but powerful architecture

- Representing local points with residual MLPs

- Permutation-invariant

- Lightweight counterpart, PointMLP-elite