

Point-LN: A Lightweight Framework for Efficient Point Cloud Classification Using Non-Parametric Positional Encoding

Marzieh Mohammadi
Sirjan University of Technology, Iran
m.mohammadi@stu.sirjantech.ac.ir

Amir Salarpour
Sirjan University of Technology, Iran
salarpour@sirjantech.ac.ir

Pedram MohajerAnsari
Clemson University, USA
pmohaje@clemson.edu

Abstract—We introduce *Point-LN*, a novel lightweight framework engineered for efficient 3D point cloud classification. *Point-LN* integrates essential non-parametric components—such as Farthest Point Sampling (FPS), k-Nearest Neighbors (k-NN), and non-learnable positional encoding—with a streamlined learnable classifier that significantly enhances classification accuracy while maintaining a minimal parameter footprint. This hybrid architecture ensures low computational costs and rapid inference speeds, making *Point-LN* ideal for real-time and resource-constrained applications. Comprehensive evaluations on benchmark datasets, including ModelNet40 and ScanObjectNN, demonstrate that *Point-LN* achieves competitive performance compared to state-of-the-art methods, all while offering exceptional efficiency. These results establish *Point-LN* as a robust and scalable solution for diverse point cloud classification tasks, highlighting its potential for widespread adoption in various computer vision applications. For more details, see the code at: https://github.com/asalarpour/Point_LN.

Index Terms—3D Point Cloud Classification, Lightweight Framework, Non-Parametric Positional Encoding, Machine Learning, Computer Vision

I. INTRODUCTION

In the field of computer vision (CV), the technology for processing two-dimensional images nears maturity [1] [2], and researchers are increasingly shifting their focus toward three-dimensional scenes, which more accurately represent real-world environments. Point clouds, consisting of large collections of tiny points in 3D space, play a crucial role in this process due to their rich informational content. These unordered sets of multidimensional points become meaningful when considered together, with point cloud classification being a key task in point cloud analysis. This task is essential across various domains, including security detection [3], [4] [5], target object detection [6], three-dimensional reconstruction [7], and autonomous vehicles [8]–[10].

Point cloud data, due to its inherent irregularity and sparsity, presents significant challenges for processing and analysis. Deep learning-based models for point cloud classification have

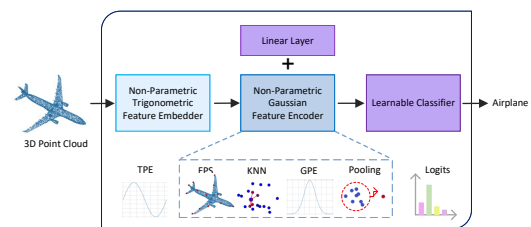


Fig. 1. Illustration of *Point-LN* for Point Cloud Classification.

gained widespread adoption due to their advantages in feature extraction and high classification accuracy. Early methods such as PointNet [11] and PointNet++ [12] introduced innovative frameworks for directly processing unordered point clouds, achieving remarkable performance. Subsequent approaches, including PointConv [13] and PointMLP [14], further advanced the field by incorporating more sophisticated feature aggregation mechanisms, leading to state-of-the-art results in various benchmarks.

Despite their success, these parametric models have notable limitations. They often require large amounts of training data, substantial computational resources, and extensive hyperparameter tuning to achieve optimal performance. Moreover, their ability to generalize can be compromised in scenarios involving diverse datasets or varying conditions, limiting their adaptability in practice.

To address these challenges, non-parametric methods such as Point-NN [15] and Point-GN [16] have emerged as promising alternatives to parametric models in point cloud analysis. These methods eliminate the need for extensive parameter optimization and training, instead relying on techniques like nearest-neighbor search and similarity-based approaches. Their simplicity, flexibility, and ability to perform inference across diverse data distributions make them particularly suitable for scalable and robust solutions in tasks such as object detection and 3D reconstruction.

However, despite their adaptability, non-parametric methods often struggle to achieve high accuracy in complex scenarios. To address these limitations, we propose a novel

hybrid framework, *Point-LN*, which combines the strengths of non-parametric techniques with a learnable classifier. This approach improves classification accuracy while maintaining reduced reliance on learnable parameters, as illustrated in Figure 1. The proposed framework enhances versatility and effectiveness, making it highly applicable to real-world tasks like object detection and 3D reconstruction.

In this paper, we introduce an efficient network for point cloud classification that effectively captures and adaptively aggregates multivariate geometric features, without relying on learnable parameters for feature extraction. Additionally, we incorporate a classifier that uses a minimal number of learnable parameters, thereby enhancing both efficiency and performance.

The primary contributions of this paper are as follows:

- We present a parameter-efficient derivative of non-parametric methods, delivering exceptional performance without relying on complicated operators.
- We incorporate a learnable classifier that effectively maps the extracted features to target categories, allowing for fine-tuning and improved classification performance.
- Despite its simple design philosophy, our approach demonstrates strong performance in 3D point cloud classification, achieving competitive results on popular benchmarks such as ModelNet40 [17] and ScanObjectNN [18].

II. RELATED WORKS

Methods for 3D point cloud classification can be broadly categorized into *Projection-based* and *Point-based* approaches. The following subsections review these two categories in detail and discuss their key techniques, advantages, and limitations.

A. Projection-based methods

Projection-based methods address the irregularity and sparsity of point clouds by transforming them into grid-like representations, such as multi-view depth maps [19] [20] and 3D voxels [21] [22]. By utilizing sets of 2D views, which contain extensive information for recognizing 3D shapes, these methods, like MVCNN [23], use rendered 2D images to identify 3D objects. Similarly, VoxNet [24] introduced a convolutional neural network (CNN) architecture that represents 3D data using a volumetric occupancy grid. While effective in processing 3D data, projection-based methods have limitations. Transforming point clouds into 2D or voxelized grids can result in the loss of fine-grained spatial relationships, potentially affecting classification performance, especially for complex tasks that require preserving precise geometric details.

B. Point-based Methods

Recent advances have focused on directly processing point clouds without transforming them into grids. *Point-based methods* can be broadly categorized into *Parametric* (deep learning-based) and *Non-parametric* methods, as follows:

1) *Parametric Methods*: They typically rely on deep learning techniques and can be subdivided into several groups:

- (i) **Convolution-based methods** [13], [25]–[27] use convolutional neural networks (CNN) to extract spatial features from point clouds. By applying convolutions over local neighborhoods of points, these methods capture spatial hierarchies and local geometric structures. This technique is highly effective for tasks such as classification, segmentation, and object recognition. For example, PointNet [11] and PointNet++ [12] employ convolutional layers to learn point-wise features, aggregating them to form a global representation of the point cloud. These models excel at handling unordered, irregular point clouds, making them suitable for a variety of 3D vision tasks.
- (ii) **Graph-based methods** [28], [29] treat point clouds as graphs, where each point is represented as a node, and edges denote relationships or distances between points. This representation is particularly effective for capturing connectivity and local structures within the point cloud. Notably, Dynamic Graph CNN (DGCNN) [29] dynamically constructs graphs during training, allowing the model to learn both local and global features. This dynamic graph construction improves the model’s ability to capture complex geometric relationships, enhancing performance on tasks like classification and segmentation.
- (iii) **Transformer-based methods** [30] [31] leverage self-attention mechanisms to process point clouds, allowing the model to weigh the importance of different points based on their spatial relationships. This ability to capture long-range dependencies is particularly useful for tasks where global context is crucial. For example, Point Transformer [31] introduces ScorNet, a learning score-based focus module.

Point Transformer first extracts local and global features and then uses ScorNet to rank these local features. A local-global attention mechanism is applied to associate the ranked local features with global features, allowing the model to focus on the most relevant features for classification and segmentation. While parametric models have demonstrated high accuracy, they often come with high computational complexity, limiting their applicability in real-time scenarios.

2) *Non-parametric Methods*: These methods, unlike parametric ones, do not require training or parameter tuning, which eliminates much of the computational overhead. These methods focus on directly capturing geometric features from point clouds, making them computationally efficient and adaptable to diverse data without the need for extensive training.

- (i) **Point-NN** [15] leverages non-parametric techniques to effectively capture geometric features by focusing on local neighborhood information. By minimizing reliance on learnable parameters, Point-NN enhances adaptability while maintaining simplicity.
- (ii) **Point-GN** [16] Similarly builds on these concepts by introducing a Gaussian embedding function that improves

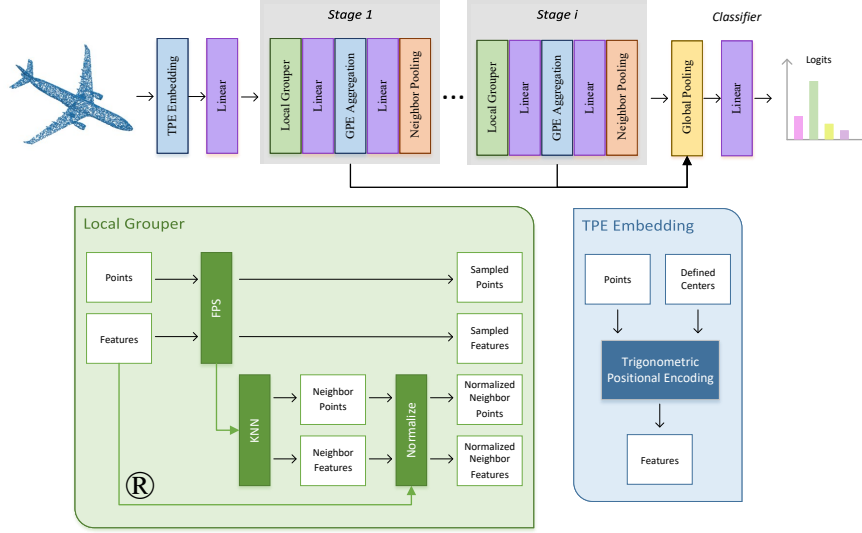


Fig. 2. Overview of the proposed network architecture: the feature encoder extracts high-dimensional representations from raw point clouds, and the classifier maps these features to the target label space.

classification accuracy without adding significant computational overhead. Despite their simplicity and flexibility, non-parametric methods often fall short of the high accuracy levels seen in parametric models, especially for complex classification tasks. This trade-off between simplicity and performance underscores the need for further advancements in the field.

- (iii) **Point-PN** [15] effectively combines the strengths of both parametric and non-parametric methods to bridge this gap. By integrating a conventional learnable classifier and enhancing the embedding with parametric linear layers, Point-PN captures higher-level spatial patterns while maintaining simplicity. Additionally, it employs trigonometric positional encoding to better represent geometric relationships within the point cloud, enabling the model to extract rich spatial features. This approach demonstrates that a powerful and efficient 3D model can be developed from a non-parametric framework without relying on complex operators or an excessive number of parameters.

III. METHOD

In this section, we first introduce the fundamentals of 3D point clouds and review common classification techniques. Next, we explore the application of non-parameteric positional encodings in lightweight methods and how they enhance the processing of point cloud data. We then outline the architecture of our proposed network **Point-LN**, which consists of two key components: the *feature encoder* and the *classifier*. The feature encoder transforms the raw input point cloud into high-dimensional feature representations, while the classifier maps these encoded features to the target label space, producing classification logits. A detailed overview of the architecture is provided in Figure 2.

A. Background

Point cloud classification aims to assign a label to an entire point cloud, or to specific regions within the point cloud, allowing for the identification of object characteristics. A 3D point cloud is a set of points $\mathcal{P} = \{p_1, p_2, \dots, p_N\}$, where each point p_i is represented by its 3D coordinates (x_i, y_i, z_i) , and optionally additional features $\gamma(p_i) \in \mathbb{R}^{C \times 3}$, where C is the feature dimension. The goal is to process this input point cloud and classify it into one of several predefined categories. The classification process consists of two main steps: the feature encoder and the classifier. First, the feature encoder extracts meaningful representations from the raw point cloud by aggregating spatial and geometric information across points. The encoded feature vector \mathbf{F} represents the output of the feature encoder, which processes the input point cloud:

$$\mathbf{F} = \text{Encoder}(\{p_i\}_{i=1}^N) \quad (1)$$

The second step, the classifier, takes \mathbf{F} as input to predict the point cloud's category, typically using fully connected layers followed by a softmax activation. The classification output \mathbf{y} is given by:

$$\mathbf{y} = \text{Softmax}(\text{Classifier}(\mathbf{F})) \quad (2)$$

B. Positional Encoding

Positional encoding (PE), first introduced in the transformer architecture [32], has proven successful in various domains such as natural language processing (NLP) and CV. For point cloud processing, positional encoding captures essential spatial relationships between points, enabling the model to understand their relative positions in 3D space. We explore two types of positional encoding commonly used in point cloud processing.

(i) Trigonometric Positional Encoding (TPE)

Point-NN [15] employs TPE, taking advantage of the periodic nature of sine and cosine functions. This method effectively captures the spatial relationships between points in 3D space. For a point $p_i = (x_i, y_i, z_i)$, trigonometric functions are used to generate a C_I -dimensional positional embedding. The encoding for each axis is defined as follows:

$$\text{PE}_{\text{TPE}} \begin{cases} \gamma_x(x_i, 2n) = \sin(\alpha x_i / \beta^{\frac{6n}{C_I}}) \\ \gamma_x(x_i, 2n+1) = \cos(\alpha x_i / \beta^{\frac{6n}{C_I}}) \\ \gamma_y(y_i, 2n) = \sin(\alpha y_i / \beta^{\frac{6n}{C_I}}) \\ \gamma_y(y_i, 2n+1) = \cos(\alpha y_i / \beta^{\frac{6n}{C_I}}) \\ \gamma_z(z_i, 2n) = \sin(\alpha z_i / \beta^{\frac{6n}{C_I}}) \\ \gamma_z(z_i, 2n+1) = \cos(\alpha z_i / \beta^{\frac{6n}{C_I}}) \end{cases} \quad (3)$$

where C_I is the initial feature dimension, α controls the scale, and β governs the wavelength. The final positional encoding for each point is given by:

$$\gamma^{\text{TPE}}(p_i) = \text{Concat}(\gamma_x(x_i); \gamma_y(y_i); \gamma_z(z_i)) \in \mathbb{R}^{1 \times C_I} \quad (4)$$

- (ii) **Gaussian Positional Encoding (GPE)** GPE, used in Point-GN [16], enhances spatial awareness by encoding positional information into the feature representation without requiring learnable parameters. GPE uses a Gaussian function to map the raw point coordinates into higher-dimensional space, preserving both local and global spatial information. For each axis, the encoding is defined as:

$$\text{PE}_{\text{GPE}} \begin{cases} \gamma_x(x_i, v_j) = \exp\left(-\frac{\|x_i^2 - v_j\|^2}{2\sigma^2}\right) \\ \gamma_y(y_i, v_j) = \exp\left(-\frac{\|y_i^2 - v_j\|^2}{2\sigma^2}\right) \\ \gamma_z(z_i, v_j) = \exp\left(-\frac{\|z_i^2 - v_j\|^2}{2\sigma^2}\right) \end{cases} \quad (5)$$

where v_j represents reference points and σ is the standard deviation that influences the scale of local versus global features. The final positional encoding for each point is the concatenation of embeddings across all axes:

$$\gamma^{\text{GPE}}(p_i) = [\gamma_x(x_i, v_j), \gamma_y(y_i, v_j), \gamma_z(z_i, v_j)]_{j=1}^V \quad (6)$$

where V is the number of reference points along each axis.

C. Lightweight Feature Encoder

We provide a lightweight feature encoder, and to better extract multi-scale hierarchy, we append simple linear layers to each stage of the encoder. In the following sections, we will outline the four steps of this feature encoder.

1) Initial Embedding

TPE. To perform feature embedding we utilize *TPE*, which transforms the Cartesian coordinates of each point $p_i \in \mathbb{R}^d$ into a trigonometric representation, preserving

the spatial relationships and helping the model maintain a consistent understanding of geometry. The resulting embeddings, $\gamma^{\text{TPE}}(p_i)$, encapsulate geometric features derived from the point cloud.

Linear Layer. The trigonometric embeddings $\gamma^{\text{TPE}}(p_i)$ are then passed through a linear layer, which performs a learnable linear transformation. This operation is defined as:

$$f_i = \text{Linear}(\gamma^{\text{TPE}}(p_i)) = \mathbf{W} \cdot \gamma^{\text{TPE}}(p_i) + \mathbf{b},$$

where \mathbf{W} is a learnable weight matrix, \mathbf{b} is a bias vector, and f_i represents the output embedded features for point p_i . This step embeds the TPE-transformed features, enabling the model to capture more complex patterns in the input data.

Final Output. The output of the initial embedding process includes the original points and their associated embedded features, represented as:

$$\{p_i, f_i\}_{i=1}^N,$$

where p_i are the original Cartesian coordinates of the points, and $f_i = \text{Linear}(\gamma^{\text{TPE}}(p_i))$ are the embedded features derived from the TPE-transformed coordinates.

2) Local Grouper

Sampling Layer. At each stage, the input point cloud and associated features from the previous stage is represented as $\{p_i, f_i\}_{i=1}^N$ the features could be TPE from initial embed step or GPE for different stages. To reduce computational complexity while retaining the global structure of the point cloud, we perform *Farthest Point Sampling (FPS)*. This step selects a subset of $N/2$ points that are maximally spaced apart, ensuring they represent the center of the point cloud:

$$\{\mathbf{p}_j, f_j\}_{j=1}^{N/2} = \text{FPS}\left(\{\mathbf{p}_i, f_i\}_{i=1}^N\right) \quad (7)$$

Grouping Layer. We then group the points into local neighborhoods using the *K-Nearest Neighbors (KNN)* algorithm. This step allows us to capture local geometric patterns:

$$\text{idx}_j = \text{KNN}(\mathbf{p}_j, \mathbf{p}_i) \quad (8)$$

where idx_j represents the indices of the K nearest neighbors for point \mathbf{p}_j . The retrieved coordinates and features are:

$$\mathbf{P}_j = \text{retrieve}\left(\{\mathbf{p}_i\}_{i=1}^N, \text{idx}_j\right) \in \mathbb{R}^{K \times 3} \quad (9)$$

$$\mathbf{F}_j = \text{retrieve}\left(\{f_i\}_{i=1}^N, \text{idx}_j\right) \in \mathbb{R}^{K \times (V \times 3)} \quad (10)$$

where, \mathbf{P}_j denotes the gathered coordinates, and \mathbf{F}_j represents the gathered features for the point \mathbf{p}_j . The retrieved coordinates \mathbf{P}_j and features \mathbf{F}_j are then normalized using the mean and standard deviation of each point's neighbors. These normalized coordinates and features

are subsequently forwarded to the next stage for further processing.

3) Local Geometry Aggregation

GPE Aggregation. The features from the Local Grouper are then fed into the *Local Geometry Aggregation*. Here, we use *GPE* to extract meaningful spatial information, that helps encode global and local geometric patterns, ensuring a rich feature representation without the need for learnable parameters. The spatially encoded features are combined with the neighborhood features through element-wise multiplication:

$$\mathbf{F}_j \leftarrow \mathbf{F}_j + \gamma(\mathbf{P}_j) \odot \gamma(\mathbf{P}_j) \quad (11)$$

Linear Layer. Here, we insert two learnable linear layers for each stage, positioned right before and after the GPE Aggregation step. This addition aims to capture higher-level spatial information. Specifically, by placing these layers strategically, we enhance the encoder's ability to process and represent geometric features effectively.

4) Pooling

Neighbor Pooling. To aggregate local features, we use both mean and max pooling, ensuring that the final features are permutation-invariant and capture local geometry from multiple perspectives:

$$\Phi_j = \text{Mean}(\mathbf{F}_j) + \text{Max}(\mathbf{F}_j), \forall j \in \mathbb{Z}_K \quad (12)$$

where $\text{Mean}(\mathbf{F}_j)$ and $\text{Max}(\mathbf{F}_j)$ are permutation-invariant operations. These operations capture neighboring features from multiple perspectives and ensure that the order of neighbors has no effect on the final pooled features.

Aggregation Across Stages. The Lightweight feature encoder includes four stages, each producing pooled features Φ_j^s . After processing through all stages, global pooling is performed on the results from each stage. The final feature vector \mathbf{F} for the lightweight feature encoder is obtained by concatenating the global mean and max features from all four stages:

$$\mathbf{F} = \bigoplus_{s=1}^4 \left[\text{Mean}(\Phi_j^s) + \text{Max}(\Phi_j^s) \right] \quad (13)$$

which captures and aggregates spatial and feature information across multiple levels by integrating the mean and max features from each stage.

D. Classifier Architecture

The classifier maps the high-dimensional feature space, produced by the lightweight encoder, directly to the target label space. This mapping is performed by a simple neural network, which generates the output logits. The predicted probabilities are then obtained using the softmax function:

$$\mathbf{y} = \text{Softmax}(\text{Classifier}(\mathbf{F})), \quad (14)$$

where $\text{Classifier}(\cdot)$ represents the transformation from the feature space \mathbf{F} to the logits.

The final predicted class is determined by selecting the class with the highest probability:

$$c = \arg \max(\mathbf{y}). \quad (15)$$

IV. EXPERIMENTS

In this section, we evaluate the performance of our method for point cloud classification through experiments conducted on two well-known datasets: ModelNet40 [17] and ScanObjectNN [18]. ModelNet40 consists of clean, synthetic 3D models, providing a controlled environment for assessing classification performance on ideal data. In contrast, ScanObjectNN presents a more challenging scenario with real-world 3D data that includes occlusions, clutter, and background noise, thus offering a more rigorous test of the robustness and versatility of our approach.

To thoroughly evaluate our method, we compare it against several state-of-the-art techniques, including both fully trained models and non-parametric methods. Specifically, we benchmark our approach against PointMLP [14], a recent fully trained model that has demonstrated impressive results in point cloud classification tasks. Additionally, we assess the performance of non-parametric methods such as Point-NN [15] and Point-GN [16], which have shown promising results in similar classification settings. Through these comprehensive comparisons, we aim to highlight the strengths of our method, demonstrating its effectiveness in handling both synthetic and real-world data, while also emphasizing its competitive edge in the context of fully trained and non-parametric models.

A. Experimental Setup

We evaluate the performance of *Point-LN* on a system equipped with an NVIDIA RTX 4090 GPU. Although *Point-LN* is designed to be lightweight, the use of the high-performance GPU significantly accelerates both the training and inference processes. This allows for efficient model evaluation and rapid experimentation, particularly when dealing with large-scale and complex 3D datasets such as *ModelNet40* [17] and *ScanObjectNN* [18].

While *Point-LN* does not require a vast amount of training resources, the GPU's computational power enables faster convergence and efficient handling of the diverse real-world data. This ensures that we can conduct comprehensive benchmarking, rapidly assess the model's performance, and explore various hyperparameter configurations. The use of this hardware ensures our method can be evaluated effectively across both synthetic and real-world datasets, highlighting its strengths in terms of efficiency and scalability.

B. Dataset Details

The *ModelNet40* [17] dataset consists of 12,311 CAD models across 40 object categories, split into 9,843 samples for training and 2,468 for testing. This dataset is widely used

TABLE I
COMPARISON OF POINT CLOUD CLASSIFICATION METHODS ON MODELNET40

Method	Acc. (%)	Param.
Point-NN [15]	81.8	0.0 M
Point-GN [16]	85.3	0.0 M
PointNet [11]	89.2	3.5 M
PointNet++ [12]	90.7	1.7 M
DGCNN [29]	92.9	1.8 M
GBNet [33]	93.8	8.4 M
CurveNet [34]	93.8	2.0 M
Point-PN [15]	93.8	0.8 M
PointMLP [14]	94.1	12.6 M
Point-LN	94.0	0.8 M

for point cloud classification due to its clean, synthetic nature, providing a controlled environment for benchmarking.

In contrast, the *ScanObjectNN* [18] dataset presents a more challenging real-world scenario, with 2,902 samples across 15 object categories. Objects in *ScanObjectNN* are often occluded, cluttered, or contain background noise, providing a closer simulation to real-world 3D data. The dataset is divided into three official subsets: *OBJ-BG*, which contains objects with background noise, *OBJ-ONLY*, with objects without background, and *PB-T50-RS*, featuring partial occlusions and transformations. These subsets test the robustness of models under various degrees of complexity. For both datasets, we follow the common practice of sampling 1,024 points from each object, as used in prior works (e.g., PointNet++ [12], DGCNN [29]). Our model combines *maximum pooling* and *average pooling* to enhance feature aggregation, inspired by DGCNN [29].

C. Shape classification on ModelNet40

Table I evaluates the performance of **Point-LN** on the *ModelNet40* [17] dataset. It achieves an accuracy of 94.0%, which is comparable to the best performing model, PointMLP [14], with an accuracy of 94.1%. However, PointMLP has a significantly larger model size (12.6M parameters). In contrast, our approach demonstrates the ability to capture both local and global geometric features while maintaining minimal model complexity, with only 0.8 million parameters. When compared to other methods, **Point-LN** shows competitive performance with models such as PointNet [11] (89.2%) and PointNet++ [12] (90.7%), while being much more lightweight.

For instance, **Point-LN** achieves a higher accuracy than Point-GN [16] (85.3%) and Point-NN [15] (81.8%), both of which have zero parameters but show lower classification accuracy. This demonstrates the effectiveness of **Point-LN** in extracting meaningful features with only a small number of parameters. Additionally, **Point-LN** ensures high efficiency for real-time applications. Despite its compact size, it provides performance on par with larger parametric models like PointMLP, but with substantially reduced computational overhead. This combination of competitive accuracy and exceptional efficiency makes **Point-LN** an ideal choice for resource-constrained environments, where both real-time performance and minimal model complexity are crucial.

D. Shape classification on ScanObjectNN

While the *ModelNet40* [17] dataset is a well-established benchmark for point cloud analysis, its synthetic nature may limit the evaluation of methods under more challenging, real-world conditions. To address this, we also, we evaluate the performance of **Point-LN** on the challenging *ScanObjectNN* [18] benchmark, presenting our results in Table II. **Point-LN** achieves state-of-the-art accuracy across all subsets, with 92.2% on *OBJ-BG*, 92.1% on *OBJ-ONLY*, and 91.7% on the most challenging *PB-T50-RS* subset. These results highlight the robustness of **Point-LN** in handling real-world 3D data with noise, occlusions, and clutter.

TABLE II
COMPARISON OF POINT CLOUD CLASSIFICATION METHODS ON SCANOBJECTNN

Method	OBJ-BG	OBJ-ONLY	PB-T50-RS	Param.
Point-NN [15]	71.1	74.9	64.9	0.0 M
Point-GN [16]	85.2	86.0	86.4	0.0 M
PointNet [11]	73.3	79.2	68.2	3.5 M
PointNet++ [12]	82.3	84.3	77.9	1.7 M
DGCNN [29]	82.8	86.2	78.1	1.8 M
PointCNN [25]	86.1	85.5	78.5	-
GBNet [33]	-	-	80.5	8.4 M
PointMLP [14]	-	-	85.4	12.6 M
Point-PN [15]	91.0	90.2	87.1	0.8 M
Point-LN	92.2	92.1	91.7	0.8 M

Compared to parametric methods, **Point-LN** delivers substantial accuracy improvements while maintaining a lightweight architecture. For instance, it outperforms PointNet [11] by +18.9% on *OBJ-BG* and PointNet++ [12] by +9.9% on *OBJ-ONLY*. Furthermore, **Point-LN** matches the performance of the larger PointMLP [14] (91.7% vs. 85.4% on *PB-T50-RS*) while requiring only 0.8 million parameters, compared to PointMLP's 12.6 million. Similar to non-parametric methods such as Point-NN [15] and Point-GN [16], **Point-LN** leverages positional encoding for feature extraction without relying heavily on trainable parameters.

However, **Point-LN** introduces a novel lightweight architecture that significantly enhances performance, achieving a +10.3% improvement over Point-GN on *PB-T50-RS* while maintaining a small parameter count. These results demonstrate the effectiveness of **Point-LN** in combining the strengths of non-parametric techniques with enhanced feature extraction capabilities. Its ability to deliver high accuracy with minimal computational overhead makes **Point-LN** highly suitable for resource-constrained environments and real-time applications.

V. CONCLUSION

In this paper, we presented **Point-LN**, a lightweight framework tailored for efficient 3D point cloud classification. By integrating essential non-parametric components—such as Farthest Point Sampling (FPS), k-Nearest Neighbors (k-NN), and non-learnable positional encoding—with a streamlined learnable classifier, **Point-LN** achieves high classification accuracy with minimal parameters. This design ensures both computational efficiency and rapid inference, making it

suitable for real-time and resource-constrained applications. Our experiments on benchmark datasets, ModelNet40 and ScanObjectNN, demonstrate that **Point-LN** matches or outperforms state-of-the-art methods while maintaining a significantly smaller model size.

The framework's robustness and adaptability are evident in its ability to handle diverse and noisy real-world 3D data effectively. Future work will explore enhancing **Point-LN** by incorporating additional geometric features and extending its application to more complex 3D tasks such as semantic segmentation and object detection. These advancements aim to further improve scalability and performance, solidifying **Point-LN**'s role as a versatile tool in 3D point cloud analysis. In summary, **Point-LN** offers a balanced solution that combines high accuracy with exceptional efficiency, positioning it as a promising framework for various computer vision applications involving 3D point clouds.

REFERENCES

- [1] L. Zhang, L. Sun, W. Li, J. Zhang, W. Cai, C. Cheng, and X. Ning, "A joint bayesian framework based on partial least squares discriminant analysis for finger vein recognition," *IEEE Sensors Journal*, vol. 22, no. 1, pp. 785–794, 2021.
- [2] X. Ning, W. Tian, Z. Yu, W. Li, X. Bai, and Y. Wang, "Hcfn: high-order coverage function neural network for image classification," *Pattern Recognition*, vol. 131, p. 108873, 2022.
- [3] X. Ning, K. Gong, W. Li, and L. Zhang, "Jwsaa: joint weak saliency and attention aware for person re-identification," *Neurocomputing*, vol. 453, pp. 801–811, 2021.
- [4] A. Salarpour, H. Khotanlou, and N. Mavridis, "Long-term estimation of human spatial interactions through multiple laser ranging sensors," in *2014 International Conference on Robotics and Emerging Allied Technologies in Engineering (iCREATE)*. IEEE, 2014, pp. 109–114.
- [5] C. Yan, G. Pang, X. Bai, C. Liu, X. Ning, L. Gu, and J. Zhou, "Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, vol. 24, pp. 1665–1677, 2021.
- [6] W. Zhangyu, Y. Guizhen, W. Xinkai, L. Haoran, and L. Da, "A camera and lidar data fusion method for railway object detection," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13 442–13 454, 2021.
- [7] B. Yang, R. Huang, J. Li, M. Tian, W. Dai, and R. Zhong, "Automated reconstruction of building lods from airborne lidar point clouds using an improved morphological scale space," *Remote Sensing*, vol. 9, no. 1, p. 14, 2016.
- [8] P. MohajerAnsari, A. Domeke, J. de Voor, A. Mitra, G. Johnson, A. Salarpour, H. Olufowobi, d. Hamad, and M. D. Pesé, "Discovering new shadow patterns for black-box attacks on lane detection of autonomous vehicles," *arXiv preprint arXiv:2409.18248*, 2024.
- [9] M. Aldeen, P. MohajerAnsari, J. Ma, M. Chowdhury, L. Cheng, and M. D. Pesé, "Wip: A first look at employing large multimodal models against autonomous vehicle attacks," in *ISOC Symposium on Vehicle Security and Privacy (VehicleSec '24)*, 2024.
- [10] M. Aldeen, P. MohajerAnsari, J. Ma, M. Chowdhury, L. Cheng, and M. D. Pesé, "An initial exploration of employing large multimodal models in defending against autonomous vehicles attacks," in *2024 IEEE Intelligent Vehicles Symposium (IV)*, 2024, pp. 3334–3341.
- [11] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.
- [12] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," *Advances in neural information processing systems*, vol. 30, 2017.
- [13] W. Wu, Z. Qi, and L. Fuxin, "Pointconv: Deep convolutional networks on 3d point clouds," in *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, 2019, pp. 9621–9630.
- [14] X. Ma, C. Qin, H. You, H. Ran, and Y. Fu, "Rethinking network design and local geometry in point cloud: A simple residual mlp framework," *arXiv preprint arXiv:2202.07123*, 2022.
- [15] R. Zhang, L. Wang, Z. Guo, Y. Wang, P. Gao, H. Li, and J. Shi, "Parameter is not all you need: Starting from non-parametric networks for 3d point cloud analysis," *arXiv preprint arXiv:2303.08134*, 2023.
- [16] M. Mohammadi and A. Salarpour, "Point-gn: A non-parametric network using gaussian positional encoding for point cloud classification," *arXiv preprint arXiv:2412.03056*, 2024.
- [17] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, "3d shapenets: A deep representation for volumetric shapes," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1912–1920.
- [18] M. A. Uy, Q.-H. Pham, B.-S. Hua, T. Nguyen, and S.-K. Yeung, "Re-visiting point cloud classification: A new benchmark dataset and classification model on real-world data," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1588–1597.
- [19] S. Bai, X. Bai, Z. Zhou, Z. Zhang, and L. Jan Latecki, "Gift: A real-time and scalable 3d shape search engine," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5023–5032.
- [20] A. Hamdi, S. Giancola, and B. Ghanem, "Mvtn: Multi-view transformation network for 3d shape recognition," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 1–11.
- [21] Y. Zhou and O. Tuzel, "Voxelnet: End-to-end learning for point cloud based 3d object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4490–4499.
- [22] Y. Li, S. Pirk, H. Su, C. R. Qi, and L. J. Guibas, "Fpnn: Field probing neural networks for 3d data," *Advances in neural information processing systems*, vol. 29, 2016.
- [23] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller, "Multi-view convolutional neural networks for 3d shape recognition," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 945–953.
- [24] D. Maturana and S. Scherer, "Voxnet: A 3d convolutional neural network for real-time object recognition," in *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2015, pp. 922–928.
- [25] Y. Li, R. Bu, M. Sun, W. Wu, X. Di, and B. Chen, "Pointcnn: Convolutional on x-transformed points," *Advances in neural information processing systems*, vol. 31, 2018.
- [26] H. Thomas, C. R. Qi, J.-E. Deschaud, B. Marcotegui, F. Goulette, and L. J. Guibas, "Kpconv: Flexible and deformable convolution for point clouds," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6411–6420.
- [27] S. M. Peyghambarzadeh, F. Azizmalayeri, H. Khotanlou, and A. Salarpour, "Point-planenet: Plane kernel based convolutional neural network for point clouds analysis," *Digital Signal Processing*, vol. 98, p. 102633, 2020.
- [28] Q. Lu, C. Chen, W. Xie, and Y. Luo, "Pointngcnn: Deep convolutional networks on 3d point clouds with neighborhood graph filters," *Computers & Graphics*, vol. 86, pp. 42–51, 2020.
- [29] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph cnn for learning on point clouds," *ACM Transactions on Graphics (tog)*, vol. 38, no. 5, pp. 1–12, 2019.
- [30] J. Yang, Q. Zhang, B. Ni, L. Li, J. Liu, M. Zhou, and Q. Tian, "Modeling point clouds with self-attention and gumbel subset sampling," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 3323–3332.
- [31] N. Engel, V. Belagiannis, and K. Dietmayer, "Point transformer," *IEEE access*, vol. 9, pp. 134 826–134 840, 2021.
- [32] A. Vaswani, "Attention is all you need," *Advances in Neural Information Processing Systems*, 2017.
- [33] S. Qiu, S. Anwar, and N. Barnes, "Geometric back-projection network for point cloud classification," *IEEE Transactions on Multimedia*, vol. 24, pp. 1943–1955, 2021.
- [34] T. Xiang, C. Zhang, Y. Song, J. Yu, and W. Cai, "Walk in the cloud: Learning curves for point clouds shape analysis," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 915–924.