

3D Scalable Quantum Convolutional Neural Networks for Point Cloud Data Processing in Classification Applications

Hankyul Baek, Won Joon Yun, Joongheon Kim

Department of Electrical and Computer Engineering, Korea University, Seoul 02841, Republic of Korea
{67back, ywjoon95, joongheon}@korea.ac.kr

Abstract

As the noisy intermediate-scale quantum (NISQ) era has begun, a quantum neural network (QNN) is definitely a promising solution to many problems that classical neural networks cannot solve. In addition, a quantum convolutional neural network (QCNN) is now receiving a lot of attention because it can process high dimensional inputs comparing to QNN. However, due to the nature of quantum computing, it is difficult to scale up the QCNN to extract a sufficient number of features due to barren plateaus. This is especially challenging in classification operations with high-dimensional data input. However, due to the nature of quantum computing, it is difficult to scale up the QCNN to extract a sufficient number of features due to barren plateaus. This is especially challenging in classification operations with high dimensional data input. Motivated by this, a novel 3D scalable QCNN (sQCNN-3D) is proposed for point cloud data processing in classification applications. Furthermore, reverse fidelity training (RF-Train) is additionally considered on top of sQCNN-3D for diversifying features with a limited number of qubits using the fidelity of quantum computing. Our data-intensive performance evaluation verifies that the proposed algorithm achieves desired performance.

Introduction

A point cloud is expected to be desirable three-dimensional sensory data and it is widely and actively used in various fields, including robotics (François, Francis, and Roland 2015) and autonomous vehicles (Yaodong et al. 2022; Yun et al. 2022b,c). In addition, the point clouds are alternative data to precisely analyze and rebuild our surrounding real-world situations (Xiangyu et al. 2018). Currently, millions of point cloud vertices per second can be produced by laser imaging detection and ranging (LiDAR) sensors. Furthermore, according to *Global Market Insights* (Preeti and Smriti 2021), the market size of *point cloud processing related media services* will reach more than 15 billion dollars in 2030.

Spurred by the advance of LiDAR, recent studies on point cloud processing have shown that neural network (NN) can be an alternative solution for various tasks such as classification (Hackel et al. 2017), object detection (Zhou and Tuzel

2018; Guo et al. 2020), and segmentation (Landrieu and Simonovsky 2018). However, due to abundant spatial geometric information and the massive size of the point cloud, operating the point cloud processing is still challenging (Qin et al. 2019). A point cloud usually has around 100k vertices, and each vertex consists of a lot of information, *i.e.*, color and Cartesian coordinates. Therefore, classical computing methodologies are obviously harsh to utilize the point cloud in real-time and even jammed when they utilize NN.

Quantum computing can be a reasonable solution to these challenges. The quantum computing utilizes a quantum bit (qubit), the counterpart in quantum computing to the binary bit of classical computing. Due to the characteristic of qubit that ranges from 0 to 1, not exact integer 0 or 1 in classical computing, quantum computing can enlarge its computation capability on an exponential scale (Bai et al. 2022; Yun, Park, and Kim 2022; Park et al. 2021; Kong 2022). In addition, parallelization is one of the beauty of quantum computing. Thus, quantum algorithms have shown that they are able to solve several NP-hard problems in polynomial time, such as Shor algorithm (Shor 1994) and Grover search algorithm (Grover 1996), those are physically impossible in classical approaches (Yun et al. 2022a). Therefore, quantum computing can outperform classical algorithms in terms of processing speed for some problems under specific conditions (Nikoloska and Simeone 2022; Xiong et al. 2022).

In this paper, we focus on grafting quantum machine learning (QML) methodologies to point cloud data processing which requires tremendous computational costs (Liu et al. 2019). In naïve and straightforward approach to process point cloud data, increasing the number of qubits or gates can be intuitively considered. However, the trainability issue occurs because the number of local minima (*i.e.*, barren plateau) is proportion to the exponential number of gates in QNN (McClean et al. 2018). Therefore, this approach is not considerable. Fortunately, it has been proven that the barren plateaus can be eliminated in quantum convolutional neural network (QCNN) (Pesah et al. 2020). Furthermore, image processing with QCNN shows better feasibility when QCNN is used together with classical NN, *i.e.*, fully connected network (FCN) (Oh, Choi, and Kim 2020), which can be called a quantum-classical hybrid classifier. It has been experimentally proven that hybrid QML achieves considerable performance in classification tasks (Schuld and

Killoran 2019). Thus, it is better to use hybrid QCNNS in point cloud processing applications.

On top of these current research progresses, there still remain important questions on QCNN methodologies, i.e., (i) *how to upload tremendous classical data into QCNN?*, (ii) *how to train QCNN efficiently?*, and (iii) *how to make the complexity of QCNN be proportioned to performance?* In order to answer these questions, we aim to extend a 3D voxelized version of scalable QCNN under the consideration of barren plateaus, data uploading issues, and efficient training methods. First of all, we propose a 3D voxelized version of scalable QCNN, i.e., 3D scalable QCNN (sQCNN-3D). Moreover, we alleviate the data uploading issue by adopting data-reuploading (Pérez-Salinas et al. 2020). Furthermore, a new scaling strategy is proposed for leveraging the various sizes of quantum convolutional filters. Finally, in order to realize efficient training, we propose a 3-dimensional reverse fidelity-train (3D RF-Train), which let sQCNN-3D fully utilize the point cloud’s intrinsic features while utilizing only a limited number of qubits. Our proposed methods can not only resolve the aforementioned challenges but also answer the questions mentioned above.

Contributions. The major contributions of the research results in this paper can be categorized as follows.

- First of all, a novel scalable QCNN architecture for extensible 3D data processing, i.e., sQCNN-3D, in order to achieve scalability while pursuing quantum supremacy and also avoiding barren plateaus.
- Moreover, an additional sQCNN-specific training algorithm, i.e., *RF-Train*, in order to extract the intrinsic features with a finite number of qubits.
- Furthermore, a new scaling strategy is also proposed which unleashes the potential of sQCNN-3D. More filters are corroborated in order to induce higher accuracy.
- Lastly, data-intensive experiments are conducted to corroborate the superiority of sQCNN with RF-Train, in ModelNet and ShapeNet, widely used in the literature.

Organization. The rest of this paper is organized as follows. First of all, we briefly introduce the differences between CNN and QCNN; and the concept of quanvolution. After that, we present our proposed 3D scalable QCNN for point cloud processing in classification applications. Lastly, we evaluate the performance of the proposed algorithm with extensive experiments.

Preliminaries

Point Cloud Processing with Classical CNN

There are two main categories in point cloud processing, i.e., point-wise processing (Qi et al. 2017) and voxel-based processing (Wu et al. 2015; Maturana and Scherer 2015), and the later one is the major trend in the literature. A classical CNN for the voxel-based processing is mainly composed of four procedures, i.e., embedding, convolution, pooling, and prediction.

Embedding. In contrast to a 2D image, each point cloud in 3D data has a huge number of vertices. Thus, a set of vertices in each point cloud is embedded as features to reduce

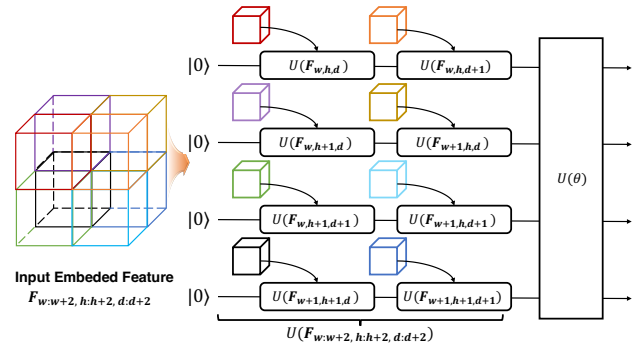


Figure 1: An illustration of multi-qubit reuploading in sQCNN-3D ($q=4, \kappa=2$).

the computational complexity as well as improve the robustness for perturbation. PointGrid (Le and Duan 2018) embeds each set of vertices into voxel grids to achieve sophisticated local geometric features; and DGCNN (Wang et al. 2019) extracts edge features from a set of vertices to incorporate local neighborhood information.

Convolution. After embedding the high dimensional point cloud data into features, the features can be convoluted by a set of filters. With trainable parameters, filters slid across each axis of features, i.e., named *input features*. The dot products between the input features and filters are computed at every spatial position.

Pooling. This pooling computation is conducted to reduce the dimension of the convoluted input features. This is usually considered as a critical part of point cloud CNN-based models because it speeds up the subsequent convolution layer computation (Wang et al. 2016). In addition, it allows these models to learn representations invariant to minor translations.

Prediction. For the prediction, the fully connected layers receive feature information which is derived from convolution layers and pooling layers. According to the universal approximation theorem (Kumagai and Sannai 2020), fully connected layers are allowed to make a prediction. By optimizing the objective function (e.g., cross-entropy loss or mean-squared error loss), conventional classical CNN can achieve the prediction to the desired classes.

Quantum CNN

QCNN is the quantum version of CNN, which leverages a parameterized quantum circuit (PQC) as convolutional filters. With the QCNN which utilizes *quanvolutional* filters (Henderson et al. 2020), spatial information can be exploited with particular characteristics. Note the definition of quanvolution is explained later.

Basic Quantum Operation. All quantum-related notations and their operations are represented as Dirac-notation. In contrast to classical computing, a qubit can have two possible states denoted as $|0\rangle$ and $|1\rangle$ (Guan et al. 2021). The quantum state in an q -qubit system is defined as 2^q possible bases and their probability amplitudes, i.e., $|\psi\rangle \triangleq \sum_{k=1}^{2^q} \alpha_k |k\rangle$ where $|k\rangle$ denotes a basis in Hilbert space,

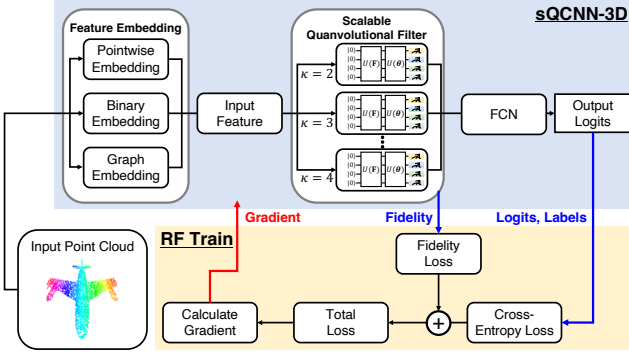


Figure 2: The pipeline for processing point clouds with sQCNN-3D.

$\forall q \in \mathbb{N}[1, \infty)$ and $\sum_{k=1}^{2^q} |\alpha_k|^2 = 1$. The operation of quantum state $|\psi\rangle$ is represented with unitary matrix U , and its operation is expressed as $|\psi\rangle \leftarrow U \cdot |\psi\rangle$. Note that the quantum state is not deterministic. Therefore, the quantum state is transformed into classical data only with its measurement; and the measurement of the quantum state is represented as a set of projection matrices $\mathcal{M} \triangleq \{\mathbf{M}_k\}_{k=1}^q$. This paper uses the measurement matrix $\mathbf{M}_k = \mathbf{I}^{\otimes k-1} \otimes \mathbf{Z} \otimes \mathbf{I}^{\otimes q-k}$, $\forall k \in [1, q]$, where \mathbf{I} denotes the identity matrix and $\mathbf{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Then, the classical output (called observable) is obtained as follows, $\langle O_k \rangle = \langle \psi | \mathbf{M}_k | \psi \rangle$. Here, the measurement is an activation function. In other words, the unitary matrix and measurement operation are mapped to linear operation and activation function, respectively, which leads to QML feasible (Hamamura and Imamichi 2020).

Quanvolution. The quanvolutional filter is defined to follow the two consecutive procedures, i.e., (i) data encoding-processing and (ii) measurement.

1. *Data Encoding-Processing:* To use quanvolutional filter with classical data, the encoding strategy should be considered where classical data is transformed into quantum states. In QML research, several encoding strategies are discussed such as basis-encoding, amplitude-encoding, and angle-encoding (Schuld and Killoran 2019). Among them, it is experimentally known that the angle-encoding presents the best performance in QML applications. Especially, the angle-encoding with data reuploading shows the high performance when the classical data \mathbf{x} are encoded to quantum states, which makes the quantum states efficiently utilize the Hilbert space using trainable parameters θ (Pérez-Salinas et al. 2020). This process can be expressed as follows,

$$|\psi\rangle = \prod_{l=0}^{\lceil \text{size}(\mathbf{x})/q \rceil} U(\theta_m) U(\mathbf{x}_{q \cdot l : q \cdot (l+1)}) |0\rangle^{\otimes q}, \quad (1)$$

where $U(\cdot)$ and $\text{size}(\cdot)$ stand for the unitary operation and the vector size of input, respectively. Because the number of variables in input data is larger than the number of qubits q in a modern noisy intermediate-scale quantum

(NISQ) era, the input data \mathbf{x} are splitted into $\mathbf{x}_{q \cdot l : q \cdot (l+1)}$ with the interval of q , $\forall l \in [1, \lceil \text{size}(\mathbf{x})/q \rceil]$. Note that $\mathbf{x}_{q \cdot l : q \cdot (l+1)}$ and θ_l denote the input vector which is composed by the first $q \cdot l$ to $q \cdot (l+1)$ elements and the trainable parameters where $\forall \theta_l \in \theta$. Similarly, the encoded quantum states are processed with PQC.

2. *Measurement:* The measurement operation has the role of not only pooling but a representation of spatial feature information, which is composed of various channels (Henderson et al. 2020). In this paper, we encode classical data to quantum states by using this angle encoding with a data reuploading strategy. In addition, we regard the measurement operation as pooling and spatial representation.

3D Scalable Quantum Convolutional Neural Networks (sQCNN-3D) for Point Cloud Processing

This section presents the motivation, architecture, training algorithm (i.e., RF-Train), and overall algorithmic procedure of our proposed sQCNN-3D, respectively.

Motivation

This paper designs sQCNN-3D for point cloud data processing in classification applications, based on following motivations. First of all, classifying point cloud data requires high computational complexity because typical point clouds in many applications contain approximately 100k points and each vertex in each point cloud uses 15 bytes (Zhou and Tuzel 2018). Second, various variational quantum algorithms suffer from vanishing gradients due to many qubits utilization. As discussed in (McClean et al. 2018), the variance of the derivative vanishes exponentially as the number of qubits increases. Therefore, it is obvious that leveraging a quantum circuit while using only a few qubits is essential during training point cloud classifiers.

sQCNN-3D Architecture

The architecture of sQCNN-3D is as shown in Fig. 2 (blue box) where it consists of two major components as follows.

Feature Embedding. We consider a point cloud that spans in a 3D Euclidean space, which is defined as $\mathcal{P} \triangleq \{(x_i, y_i, z_i) | i = 1, \dots, |\mathcal{P}|\}$, and its i -th element (i.e., vertex) is denoted as p_i . We subdivide the 3D space into voxels $\mathcal{V} \triangleq \{(\hat{x}_j, \hat{y}_j, \hat{z}_j) | j = 1, \dots, |\mathcal{V}|\}$, where its j -th voxel is denoted as v_j . Because this paper only considers the geometric features of point cloud \mathcal{P} in voxels \mathcal{V} , we obtain a geometric feature set which is denoted as $\mathcal{F} = g(\mathcal{P} \odot \mathcal{V}) \equiv \{f(v_j) | j = 1, \dots, |\mathcal{V}|, \forall f(v_j) \in \{0, 1\}\}$. Here, the operator \odot denotes that each p_i is in the area of voxel v_j . The geometric feature extraction function $g(\cdot)$ is defined according to the target geometric feature (Hinks et al. 2013), i.e., Euclidean distance, structural similarity, and binary encoding (Hinks et al. 2013). We set the range of \mathcal{P} and \mathcal{V} as $(W, H, D) \in \mathbb{R}^3$ and $(\hat{W}, \hat{H}, \hat{D}) \in \mathbb{N}^3$. From the binary encoding in (Hinks et al. 2013), $f(v_j) = 1$ if $n_j \geq T$ (otherwise, $f(v_j) = 0$) where n_j is the number of vertices that

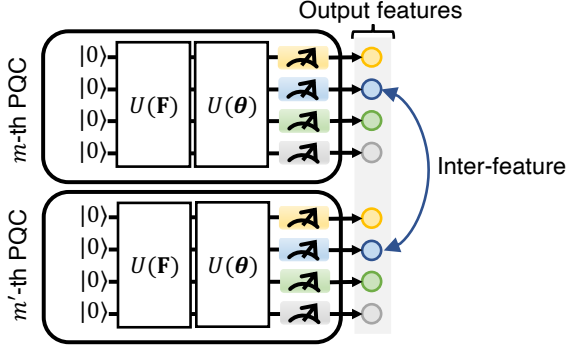


Figure 3: The meaning of terminology in RF-Train.

satisfies $\sqrt{|x_i \cdot \frac{\hat{W}}{W} - \hat{x}_j|^2} \leq 0.5$, $\sqrt{|y_i \cdot \frac{\hat{H}}{H} - \hat{y}_j|^2} \leq 0.5$ and $\sqrt{|z_i \cdot \frac{\hat{D}}{D} - \hat{z}_j|^2} \leq 0.5, \forall i \in \mathbb{N}[1, |\mathcal{P}|], \forall j \in \mathbb{N}[1, |\mathcal{V}|]$. Note, T is a threshold to determine whether v_j is active or not.

3D Data Reuploading. Inspired by (Pérez-Salinas et al. 2020), we aim to embed classical bits into quantum states using only a limited number of qubits. In contrast to classical QCNN that embeds each data into a qubit respectively, sQCNN-3D solves the huge dimensionality problem in point cloud operation through data reuploading. Fig. 1 shows the data reuploading process of sQCNN-3D. We define the input embedded feature \mathbf{F} which has the size of $c_{in} \times \hat{W} \times \hat{H} \times \hat{D}$. We interchangeably use the notation $\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}$ to represent the patch of which origin and the kernel size are (w, h, d) and κ , respectively. Note that $\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}$ is a vector with the size of κ^3 . We encode the patch $\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}$ with (1) where $\forall w \in \mathbb{N}[1, \hat{W}], \forall h \in \mathbb{N}[1, \hat{H}], d \in \mathbb{N}[1, \hat{D}]$.

Scalable 3D Quanvolutional Filter. Inspired by (Baek, Yun, and Kim 2022), we aim to design the 3D version of scalable quanvolutional filters. Fig. 2 illustrates the architecture of 3D scalable filters. In contrast to QCNN, which has limitations in extracting numerous features due to barren plateaus, sQCNN-3D scales up the filter by increasing the number of filters (*note that it does not increasing the number of qubits*). By feed-forwarding each PQC architecture, the 3D quanvolutional filter transforms $\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}$ to an observable $\langle O_{w, h, d, c} \rangle$ from a 3D matrix of scalars, where c denotes the output channel index, $\forall c \in \mathbb{N}[1, c_{out}]$. In order to extract various feature channels from the input patch, our proposed quanvolutional filter mainly consists of multiple layers of controlled-U (CU) gates. The operations of the CU gate in two qubits can be expressed as $\begin{bmatrix} I & 0 \\ 0 & U \end{bmatrix}$,

where U is illustrated as $U = \begin{bmatrix} u_{00} & u_{01} \\ u_{10} & u_{11} \end{bmatrix}$. After processing with quanvolutional filter, akin to classical CNN, sQCNN-3D achieves a feature corresponding to the measurement value of each qubit in filters. However, the scalability is not straight-forwarded yet. Thus, the quanvolution filters are

Algorithm 1: sQCNN-3D Quanvolution Procedure.

- 1 **Notation.** Input point cloud \mathcal{P} , voxel \mathcal{V} input feature \mathcal{F} and filter size κ ;
 - 2 **Input:** Input point cloud P ;
 - 3 **Initialize.** $\mathcal{F} \leftarrow g(\mathcal{P} \odot \mathcal{V})$;
 - 4 **for** $w \in \{1, 2, \dots, \hat{W}\}$ **do**
 - 5 **for** $h \in \{1, 2, \dots, \hat{H}\}$ **do**
 - 6 **for** $d \in \{1, 2, \dots, \hat{D}\}$ **do**
 - 7 Initialize quantum state, $|\psi\rangle \leftarrow |0\rangle$;
 - 8 **for** $c \in \{1, 2, \dots, c_m\}$ **do**
 - 9 Prepare data, $\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}$;
 - 10 Upload data,
 - 11 $|\psi\rangle \leftarrow U(\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}) \cdot |0\rangle^{\otimes q}$;
 - 12 Quanvolve data, $|\psi\rangle \leftarrow U(\theta) \cdot |\psi\rangle$;
 - 13 **for** $c \in \{1, 2, \dots, c_{out}\}$ **do**
 - 14 Measure the output,
 - 15 $\langle O \rangle = \langle \psi | \mathbf{I}^{\otimes n-1} \otimes \mathbf{Z} | \psi \rangle$;
 - 16 **Output:** Extracted features
-

compared with the various filter-sizes κ . This strategy enables sQCNN-3D to achieve various spatial features, which results in performance improvements.

RF-Train

In contrast to classical QCNN, sQCNN-3D can increase the number of extracted features while maintaining the number of qubits in a PQC. To achieve the full advantage of sQCNN-3D, we use RF-Train (Baek, Yun, and Kim 2022). In this paper, we consider M PQCs and each PQC is composed of q quanvolutional filters. Suppose that two PQCs exist, i.e., m -th and m' -th PQCs whose density operators are denoted as $\rho = |\psi_m\rangle\langle\psi_m|$ and $\sigma = |\psi_{m'}\rangle\langle\psi_{m'}|$, where $\forall m, m' \in \mathbb{N}[1, M]$. Denote $\langle O_k \rangle_m$ and $\langle O_k \rangle_{m'}$ as the k -th qubit's observables of m -th and m' -th PQC, respectively, where $\forall k \in [1, q]$. In quantum information theory (Bennett and Shor 1998), the similarity between $|\psi_m\rangle$ and $|\psi_{m'}\rangle$ can be measured in terms of fidelity $\Omega(\psi_m, \psi_{m'}) = (\text{Tr}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}))^2$, where $\text{Tr}(\cdot)$ is a trace operator. Note that fidelity is bounded from 0 to 1, i.e., $0 \leq \Omega(\psi_m, \psi_{m'}) \leq 1$ and $\Omega(\psi_m, \psi_m) = 1$. Briefly, fidelity between the two filters goes to 1 when the similarity between two filters increases. On the other hand, the fidelity goes to 0 when the similarity between the two filters decreases.

Motivated by the concept of fidelity, we aim to diversify inter-feature. Fig. 3 illustrates the semantic of inter-feature. It is clear that adjusting fidelity $\Omega(\psi_l, \psi_{l'})$ affects the variation of inter-features, i.e., the inter-feature distance enlarges when the fidelity between two filters is high (and vice versa). Thus, the reverse fidelity (RF) regularizer L_{RF} is designed which minimizes the fidelity between all PQCs. The regularizer is defined as follows,

$$L_{RF} = \frac{1}{M(M-1)} \sum_{m=1}^M \sum_{m' \neq m}^M \Omega(\psi_m, \psi_{m'}), \quad (2)$$

Algorithm 2: sQCNN-3D with RF-Train.

```

1 Initialization. sQCNN-3D parameters ;
2 for  $e = \{1, 2, \dots, E\}$  do
3   for  $(\mathbf{x}, y) \in \zeta$  do
4     for  $j \in [1, 2, \dots, \hat{W} \times \hat{H} \times \hat{D}]$  do
5       Get input feature  $\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}$ ;
6       for  $l, l' \in \{1, 2, \dots, L-1\}$  do
7         Proceed quanvolution in Algorithm 1;
8       Calculate the output logits and cross-entropy
9       loss with (3);
9       Calculate RF-train regularization with (2);
10      Obtain the train loss and its loss gradient;
11      Update trainable parameters of sQCNN-3D;

```

Table 1: List of simulation parameters.

Description	Value
# of filters	$\{1, 2, 4, 6, 8\}$
Size of voxels ($ \mathcal{V} $)	32^3
Optimizer	Adam
Initial learning rate	0.001
# of qubits in a quanvolutional filter	4
# of params in a scalable quanvolutional filter	48
# of params in a classical filter	64
Kernel size	$2^3, 3^3, 4^3$
Batch size	8
Test batch size	128

and note that the inter-feature becomes diverse, by minimizing L_{RF} .

Algorithm 2 represents the overall training process of sQCNN-3D with RF-Train. The set of input point clouds and their labels are denoted as (\mathbf{x}, y) . With the RF regularizer to achieve various features, we utilize cross-entropy loss L_{CE} for classification, written as

$$L_{CE} = -\frac{1}{C} \sum_{c=1}^C \log p(y_{pred} = y_c | \mathbf{x}), \quad (3)$$

where the number of classes is denoted as C . y_{pred} and y_c are predicted and actual classes. Finally, the train loss of sQCNN-3D with RF-Train is as,

$$L_{train} = \frac{1}{|\zeta|} \sum_{(\mathbf{x}, y) \in \zeta} [L_{CE} + \lambda L_{RF}], \quad (4)$$

where ζ and λ denote a batch and an RF regularizer parameter.

sQCNN-3D Overall Algorithmic Procedure

The Algorithm 1 and Algorithm 2 represent the quanvolution process and training process of sQCNN-3D, respectively. In general, it is obvious that point cloud classification tasks require high computational and memory resources because of the massive data and input size of the point clouds. In addition, many qubits are needed to encode the entire

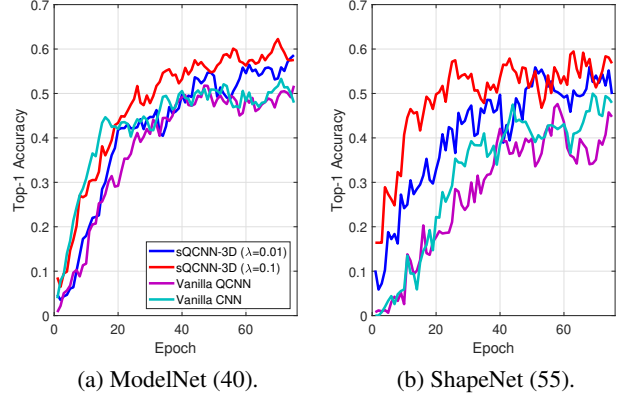


Figure 4: Top-1 accuracy of sQCNN-3D with two datasets.

point clouds into qubits, which can induce a barren plateau, in order to perform point-wise point cloud classification. As the solution under quantum computing concepts, the overall algorithmic procedure of sQCNN-3D are formalized as follows.

1. Embeds input point cloud \mathcal{P} to geometric feature set \mathcal{F} .
2. Encodes the geometric feature set to patch with origin (w, h, d) and kernel size κ . After computing the encoder, sQCNN-3D achieves $\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}$ in each stride.
3. Reuploads the input $\mathbf{F}_{w:w+\kappa, h:h+\kappa, d:d+\kappa}$ of each stride to quanvolutional filter.
4. Quanvolves the data through the quanvolutional filter.
5. Measuring and pooling occurs after quanvolution.
6. Concatenates the measured observable, i.e., extracted features, of each filter and exploit the data using FCN.

Performance Evaluation**Experimental Setting**





























To corroborate the performance of the sQCNN-3D with RF-Train, the experiments are designed as follows.

- First of all, we benchmark sQCNN-3D whether it achieves convincing accuracy as much as CNN in point cloud data classification. Our proposed sQCNN-3D with fine-tuned RF-Train is compared with classical CNN and QCNN algorithms.
- Second, we conduct an intensive experiment whether the diversity of extracted features can be increased by adjusting RF regularizer parameter λ , and it improves the performance. For this, we calculate the Euclidean distance between inter-features.
- Third, we compare the feature maps of sQCNN-3D and vanilla QCNN. Based on the visualization results as shown in Tab. 3, we corroborate that our proposed sQCNN-3D with RF-Train can fully exploit the intrinsic feature of each datum.
- Lastly, we observe the scaling strategy of sQCNN-3D. We compare sQCNN-3D with various κ to sQCNN-3D

Table 2: Top-1 accuracy and Euclidean distance with ShapeNet (55).

Metric	RF-Train		
	$\lambda = 0$	$\lambda = 0.01$	$\lambda = 0.1$
Top-1 accuracy (%)	39.5	53.14	55.3
Inter Feature distance ($\times 10^{-3}$)	0.58	0.96	2.25

Table 3: Extracted inter-feature maps according to various train strategies (i.e., untrained sQCNN-3D, Vanilla-Trained sQCNN-3D, and RF-trained sQCNN-3D).

Inputs	Untrained sQCNN-3D		sQCNN-3D with Vanilla Train		sQCNN-3D with RF-Train	
	Feature 1	Feature 2	Feature 1	Feature 2	Feature 1	Feature 2
						
						
						
						

with fixed κ . We investigate the top-1 accuracy of two strategies.

For the numerical experiments, we set the number of filters as 2 in both sQCNN and CNN models. In addition, we adopt top-1 accuracy for the metric. In addition, all convolutional filters are operated in 4 qubits system, i.e., $q = 4$. The simulation parameters are listed in Tab. 1. Note that we use the notation ‘Dataset (# of classes)’ to represent the type of dataset and the number of classes in experimental results.

Experimental Results

Performance of sQCNN-3D. Fig. 4(a)/(b) represent the point cloud classification with ModelNet40 and ShapeNet (Wu et al. 2015). In both Fig. 4(a)/(b), sQCNN-3D with RF-Train outperforms vanilla CNN and vanilla QCNN. Especially, sQCNN-3D with RF regularizer parameter $\lambda = 0.1$ achieves about 5.8% and 12.9% higher top-1 accuracy than QCNN with ModelNet40 and ShapeNet, respectively. In addition, we observe that sQCNN-3D achieves a lot higher top-1 accuracy than the CNN with 2 filters. It is the result of entanglement in quantum information theory (Bennett and Shor 1998). In contrast to CNN that conducts the dot product between the independent constituents of the input data and filter, sQCNN and QCNN entangle the input data simultaneously and it makes sQCNN and QCNN have more spatial information than CNN.

Effectiveness of RF-Train. Tab. 2 shows the inter feature distance and top-1 accuracy according to various RF regularizer parameter λ . We observe that as the RF regularizer parameter λ increases, inter feature distance increases, which leads to classification performance improvement. When the

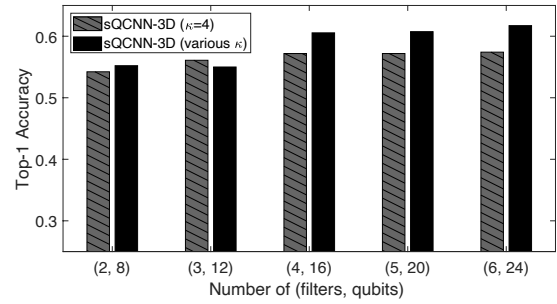


Figure 5: Scalability of sQCNN-3D with ShapeNet (55).

Table 4: Top-1 accuracy according to number of classes.

Dataset (# of classes)	Top-1 Accuracy		
	Vanilla CNN	Vanilla QCNN	sQCNN-3D
ModelNet (10)	90.3	89.73	93.7
ModelNet (20)	74.2	83.28	85.0
ModelNet (40)	46.1	53.59	59.4
ShapeNet (10)	77.6	79.8	80.9
ShapeNet (30)	63.7	66.5	68.3
ShapeNet (55)	41.8	42.4	55.3

RF regularizer parameter λ increases from 0 to 0.1, inter feature distance increases from 0.58 to 2.25 and the top-1 accuracy increases from 39.5% to 55.3%. With a limited number of qubits (features), we observed that diversity between the extracted features is significant in the classification task, and RF-Train can increase this diversity. Further to this, we visualize the impact of Vanilla-Train and RF-Train on feature diversity. Tab. 3 illustrates the individual feature maps extracted from untrained, vanilla-trained, and RF-trained filters. In contrast to the features extracted from vanilla-trained filter that show little diversity, the features extracted from RF-trained filters achieve reasonable diversity. Therefore, our proposed sQCNN-3D with RF-Train is able to exploit abundant intrinsic feature information, and this leads to performance improvements in classification tasks.

Scaling Strategy of sQCNN-3D. Fig. 5 represents the scalability of CNN and sQCNN-3D with various random filter sizes ($\forall \kappa \in \mathbb{N}[2, 4]$). Note that barren plateaus occur when the number of qubits increases in the quantum system. However, sQCNN-3D avoids the barren plateaus by increasing the number of filters while maintaining the number of qubits in a filter, even the number of (filters, qubits) increases from (2, 8) to (6, 24). In addition, we observe that combining several different kernel size κ simultaneously when scaling up the number of filters leads to performance improvements. The sQCNN-3D with various kernel sizes κ achieves 6.5% of top-1 accuracy improvement when the number of filters increases from 2 to 6, in contrast to sQCNN-3D with single kernel size $\kappa = 4$ only achieves 3.2% of top-1 accuracy improvement.

Robust Performance on the Large Number of Classes. Tab. 4 shows the top-1 accuracy according to the num-

ber of classes. sQCNN-3D achieves 93.7% top-1 accuracy on ModelNet with 10 classes. Especially, sQCNN-3D maintains the highest performance even if performance degradation occurs as the number of classes increases. In contrast to Vanilla QCNN and Vanilla CNN which show 46.8% and 37.4% of performance degradation, sQCNN-3D presents only 25.6% of performance degradation on ShapNet. In addition, Vanilla QCNN and Vanilla CNN which show 36.14% and 44.2% of performance degradation, respectively, sQCNN-3D only reveals 34.3% of performance degradation on ModelNet.

Conclusions and Future Work

This paper proposes a new 3D scalable quantum convolutional neural network (sQCNN-3D) for point cloud data classification. To fully exploit intrinsic features in point cloud data, which QCNN cannot, we design a new sQCNN-3D in order to realize scalable 3D convolutional filtering functionalities. In addition, this paper proposes RF-Train and a novel scaling strategy to maximize the training performance of sQCNN-3D. With data-intensive performance evaluation, we corroborate that our proposed sQCNN-3D outperforms conventional CNN and QCNN algorithms. We look forward to sQCNN-3D playing a leading role in point cloud data classification approaches in the near future based on its scalability and performance. Aligned with the development of quantum computing methodologies, it is obvious that the capabilities and scopes of sQCNN-3D will increase. Thus, it will be interesting to study the broad applications of sQCNN-3D-extensions based operations to large-scale point cloud data processing tasks such as object detection requiring high-accurate and high-speed data-intensive real-time computation.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00907, Development of AI Bots Collaboration Platform and Self-organizing AI). Won Joon Yun and Joongheon Kim are the corresponding authors of this paper.

References

Baek, H.; Yun, W. J.; and Kim, J. 2022. Scalable Quantum Convolutional Neural Networks. *CoRR*, abs/2209.12372.

Bai, L.; Jiao, Y.; Cui, L.; Rossi, L.; Wang, Y.; Yu, P. S.; and Hancock, E. R. 2022. Learning Graph Convolutional Networks based on Quantum Vertex Information Propagation. In *Proc. IEEE International Conference on Data Engineering (ICDE)*, 3132–3133. Kuala Lumpur, Malaysia.

Bennett, C. H.; and Shor, P. W. 1998. Quantum information theory. *IEEE Transactions on Information Theory*, 44(6): 2724–2742.

François, P.; Francis, C.; and Roland, S. 2015. A Review of Point Cloud Registration Algorithms for Mobile Robotics. *Foundations and Trends in Robotics*, 4(1): 1–104.

Grover, L. K. 1996. A Fast Quantum Mechanical Algorithm for Database Search. In *Proc. ACM Symposium on the Theory of Computing (STOC)*, 212–219. Philadelphia, PA, USA.

Guan, W.; Perdue, G. N.; Pesah, A.; Schuld, M.; Terashi, K.; Vallecorsa, S.; and Vlimant, J. 2021. Quantum machine learning in high energy physics. *Machine Learning: Science Technology*, 2(1): 11003.

Guo, Y.; Wang, H.; Hu, Q.; Liu, H.; Liu, L.; and Bennamoun, M. 2020. Deep learning for 3D point clouds: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(12): 4338–4364.

Hackel, T.; Savinov, N.; Ladicky, L.; Wegner, J. D.; Schindler, K.; and Pollefeys, M. 2017. Semantic3D.net: A new Large-scale Point Cloud Classification Benchmark. *CoRR*, abs/1704.03847.

Hamamura, I.; and Imamichi, T. 2020. Efficient evaluation of quantum observables using entangled measurements. *Quantum Information*, 6(1): 1–8.

Henderson, M.; Shakya, S.; Pradhan, S.; and Cook, T. 2020. Quantum Convolutional Neural Networks: Powering Image Recognition with Quantum Circuits. *Quantum Machine Intelligence*, 2(1): 1–9.

Hinks, T.; Carr, H.; Truong-Hong, L.; et al. 2013. Point cloud data conversion into solid models via point-based voxelization. *Journal of Surveying Engineering*, 139(2): 72–83.

Kong, P. 2022. A Review of Quantum Key Distribution Protocols in the Perspective of Smart Grid Communication Security. *IEEE Systems Journal*, 16(1): 41–54.

Kumagai, W.; and Sannai, A. 2020. Universal Approximation Theorem for Equivariant Maps by Group CNNs. *CoRR*, abs/2012.13882.

Landrieu, L.; and Simonovsky, M. 2018. Large-Scale Point Cloud Semantic Segmentation With Superpoint Graphs. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4558–4567. Salt Lake City, UT, USA.

Le, T.; and Duan, Y. 2018. PointGrid: A Deep Network for 3D Shape Understanding. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 9204–9214. Salt Lake City, UT, USA.

Liu, Z.; Tang, H.; Lin, Y.; and Han, S. 2019. Point-voxel CNN for Efficient 3D Deep Learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 963–973. Vancouver, BC, Canada.

Maturana, D.; and Scherer, S. A. 2015. VoxNet: A 3D Convolutional Neural Network for real-time object recognition. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 922–928. Hamburg, Germany.

McClellan, J. R.; Boixo, S.; Smelyanskiy, V. N.; Babbush, R.; and Neven, H. 2018. Barren plateaus in quantum neural network training landscapes. *CoRR*, abs/1803.11173.

Nikoloska, I.; and Simeone, O. 2022. Training Hybrid Classical-Quantum Classifiers via Stochastic Variational Optimization. *IEEE Signal Processing Letters*, 29: 977–981.

Oh, S.; Choi, J.; and Kim, J. 2020. A tutorial on quantum convolutional neural networks (QCNN). In *Proc. IEEE International Conference on Information and Communication Technology Convergence (ICTC)*, 236–239.

- Park, J.; Samarakoon, S.; Elgabli, A.; Kim, J.; Bennis, M.; Kim, S.-L.; and Debbah, M. 2021. Communication-Efficient and Distributed Learning Over Wireless Networks: Principles and Applications. *Proceedings of the IEEE*, 109(5): 796–819.
- Pérez-Salinas, A.; Cervera-Lierta, A.; Gil-Fuster, E.; and Latorre, J. I. 2020. Data re-uploading for a universal quantum classifier. *Quantum*, 4: 226.
- Pesah, A.; Cerezo, M.; Wang, S.; Volkoff, T.; Sornborger, A. T.; and Coles, P. J. 2020. Absence of Barren Plateaus in Quantum Convolutional Neural Networks. *CoRR*, abs/2011.02966.
- Preeti, W.; and Smriti, L. 2021. LiDAR Market Size & Growth Forecasts (2022–2030).
- Qi, C. R.; Su, H.; Mo, K.; and Guibas, L. J. 2017. PointNet: Deep learning on point sets for 3D classification and segmentation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 652–660. Honolulu, HI, USA.
- Qin, C.; You, H.; Wang, L.; Kuo, C. J.; and Fu, Y. 2019. PointDAN: A multi-scale 3D domain adaption network for point cloud representation. In *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 7190–7201. Vancouver, BC, Canada.
- Schuld, M.; and Killoran, N. 2019. Quantum Machine Learning in Feature Hilbert Spaces. *Physical Review Letters*, 122: 040504.
- Shor, P. W. 1994. Algorithms for Quantum Computation: Discrete Logarithms and Factoring. In *Proc. IEEE Symposium on Foundations of Computer Science (FOCS)*, 124–134. Santa Fe, NM, USA.
- Wang, J.; Yang, Y.; Mao, J.; Huang, Z.; Huang, C.; and Xu, W. 2016. CNN-RNN: A Unified Framework for Multi-label Image Classification. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2285–2294. Las Vegas, NV, USA.
- Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S. E.; Bronstein, M. M.; and Solomon, J. M. 2019. Dynamic Graph CNN for Learning on Point Clouds. *ACM Transactions on Graphics*, 38(5): 146:1–146:12.
- Wu, Z.; Song, S.; Khosla, A.; Yu, F.; Zhang, L.; Tang, X.; and Xiao, J. 2015. 3D ShapeNets: A deep representation for volumetric shapes. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition, (CVPR)*, 1912–1920. Boston, MA, USA.
- Xiangyu, Y.; Bichen, W.; Sanjit, S., A.; Kurt, K.; and Alberto, S., L. 2018. A LiDAR Point Cloud Generator: from a Virtual World to Autonomous Driving. In *Proc. ACM International Conference on Multimedia Retrieval (ICMR)*, 458–464. Yokohama, Japan.
- Xiong, Y.; Ng, S. X.; Long, G.; and Hanzo, L. 2022. Dual-Frequency Quantum Phase Estimation Mitigates the Spectral Leakage of Quantum Algorithms. *IEEE Signal Processing Letters*, 29: 1222–1226.
- Yaodong, C.; Ren, C.; Wenbo, C.; Long, C.; Daxin, T.; Ying, L.; and Dongpu, C. 2022. Deep Learning for Image and Point Cloud Fusion in Autonomous Driving: A Review. *IEEE Transactions on Intelligent Transportation Systems*, 23(2): 722–739.
- Yun, W. J.; Kwak, Y.; Kim, J. P.; Cho, H.; Jung, S.; Park, J.; and Kim, J. 2022a. Quantum Multi-Agent Reinforcement Learning via Variational Quantum Circuit Design. In *Proc. IEEE International Conference on Distributed Computing Systems (ICDCS)*. Bologna, Italy.
- Yun, W. J.; Park, J.; and Kim, J. 2022. Quantum Multi-Agent Meta Reinforcement Learning. *CoRR*, abs/2208.11510.
- Yun, W. J.; Park, S.; Kim, J.; and Mohaisen, D. 2022b. Self-Configurable Stabilized Real-Time Detection Learning for Autonomous Driving Applications. *IEEE Transactions on Intelligent Transportation Systems*, 1–6.
- Yun, W. J.; Shin, M.; Jung, S.; Kwon, S.; and Kim, J. 2022c. Parallelized and randomized adversarial imitation learning for safety-critical self-driving vehicles. *IEEE/KICS Journal of Communications and Networks*, 1–12.
- Zhou, Y.; and Tuzel, O. 2018. VoxelNet: End-to-end learning for point cloud based 3D object detection. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4490–4499. Salt Lake City, UT, USA.