PAPER • OPEN ACCESS

Federated quanvolutional neural network: a new paradigm for collaborative quantum learning

View the article online for updates and enhancements.

PAPER

Federated quanvolutional neural network: a new paradigm for collaborative quantum learning

Amandeep Singh Bhatia[1] , Sabre Kais[1,2] and Muhammad Ashraful Alam[1,*]

[1] School of Electrical and Computer Engineering, Purdue university, West Lafayette, IN, United States of America
[2] Department of Chemistry, Purdue University, West Lafayette, IN, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: alam@purdue.edu

Abstract

In recent years, the concept of federated machine learning has been actively driven by scientists to ease the privacy concerns of data owners. Currently, the combination of machine learning and quantum computing technologies is a hot industry topic and is positioned to be a major disruptor. It has become an effective new tool for reshaping several industries ranging from healthcare to finance. Data sharing poses a significant hurdle for large-scale machine learning in numerous industries. It is a natural goal to study the advanced quantum computing ecosystem, which will be comprised of heterogeneous federated resources. In this work, the problem of data governance and privacy is handled by developing a quantum federated learning approach, that can be efficiently executed on quantum hardware in the noisy intermediate-scale quantum era. We present the federated hybrid quantum–classical algorithm called a quanvolutional neural network with distributed training on different sites without exchanging data. The hybrid algorithm requires small quantum circuits to produce meaningful features for image classification tasks, which makes it ideal for near-term quantum computing. The primary goal of this work is to evaluate the potential benefits of hybrid quantum–classical and classical-quantum convolutional neural networks on non-independently and non-identically partitioned (Non-IID) and real-world data partitioned datasets among several healthcare institutions/clients. We investigated the performance of a collaborative quanvolutional neural network on two medical machine learning datasets, COVID-19 and MedNIST. Extensive experiments are carried out to validate the robustness and feasibility of the proposed quantum federated learning framework. Our findings demonstrate a decrease of 2%–39% times in necessary communication rounds compared to the federated stochastic gradient descent approach. The hybrid federated framework maintained a high classification testing accuracy and generalizability, even in scenarios where the medical data is unevenly distributed among clients.

1. Introduction & motivation

In the early 2010s, the era of big data facilitated the use of scientific machine learning models in multiple fields and industries. The use of deep learning techniques has incomparable benefits for several areas ranging from healthcare to finance. These areas have a high volume of sensitive data and the current centralized machine learning approach is not effective for data analytics [1]. The standard setting in machine learning considers a centralized dataset processed in a tightly integrated system. But, in the real world, data is often decentralized across many parties. The rapid advancements in machine learning have compelled individuals and organizations to rely more on data to solve decision and inference problems. But, the balanced

protection of confidentiality, integrity, and availability of data, has become a major concern beyond data security considerations.

In recent years, the concept of federated machine learning has got a significant response from industry-based and academic scholars since it was introduced by Google in 2017 [2]. The main objective of federated machine learning is to solve data-silos challenges in real-world applications. It is a privacy-preserving decentralized approach, which trains a global model collaboratively on raw data distributed across devices without sharing their private data. Then, each device sends its parameters to the global server at each round of communication. Further, the server performs an aggregation to build a global model that is representative of the data from all the devices and sends the updated model weights for further training. The communication rounds between the global server and distributed devices are repeated till achieving the intended accuracy. The overall objective is to develop a communication-efficient federated learning algorithm, which takes fewer communication rounds between clients and the server to improve the accuracy of the global model while ensuring data privacy or secrecy. The training and testing can be performed on smartphones, tablets, personal computer systems, and also edge devices that produce the data [3].
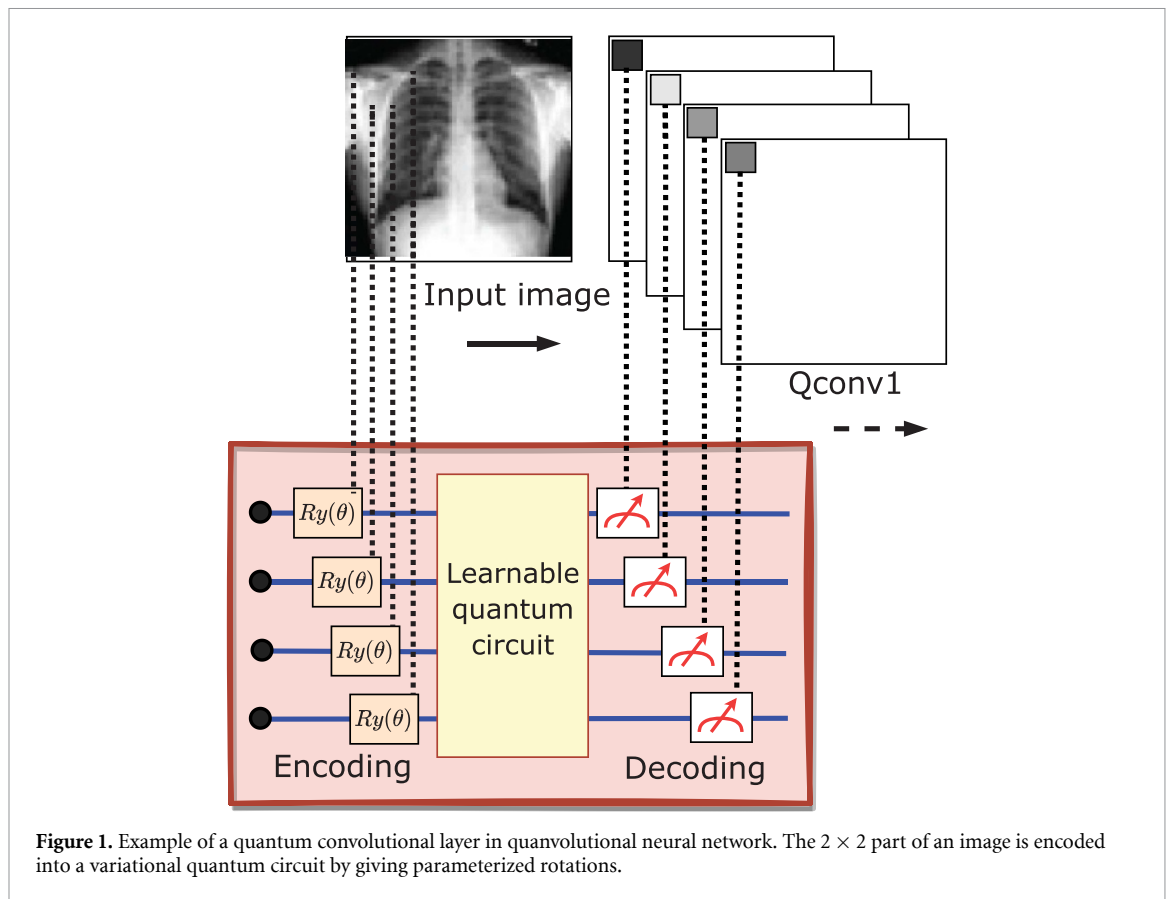
Since the introduction of federated learning based GBoard application, it has emerged as a research paradigm of future intelligent applications in several domains for collaborative learning such as smart energy infrastructures [4], healthcare [5], genomic research [6], wireless communications [7], geospatial applications [8], and many more [9]. It has shown the potential to overcome the problems of data privacy, bandwidth loss, and legalization by training algorithms collaboratively on large, distributed datasets without exchanging the data and breaching restrictions on property and privacy. Although, federated learning also poses several challenges such as quality of data, systems and statistical heterogeneity, and various kinds of attacks that can compromise either the model or the data.

In recent years, the VQCs have got an incredible response and are widely applied to several experimental schemes for noisy intermediate-scale quantum (NISQ) devices [10]. It has triggered an extensive amount of work to improve the performance of classical machine learning algorithms. It can be witnessed by the rapid increase in highly impactful quantum machine learning research papers [11–15]. It is emerging as the most exciting promising application of quantum technologies to tackle problems with improved performance using quantum phenomena entanglement and superposition. Since the introduction of quantum neural network [16, 17], many researchers focused on the quantum machine learning algorithms that have an instant influence on real-time applications, e.g. quantum chemistry [18–20] or optimization [21] to leverage the power of quantum computers.

In the NISQ era, quantum computers with 50–100 qubits, limited quantum circuit depth, and high error rates restrict the execution of several quantum gates [22]. Due to this reason, the researchers focused on hybrid algorithms to solve the computational tasks by splitting between classical and quantum computers. Several hybrid quantum–classical algorithms have been proposed for solving machine learning tasks on NISQ systems and witnessed a good performance. Recently, Henderson *et al* [23] introduced a quantum variant of convolutional neural networks (CNNs) i.e. quanvolutional neural networks (QCNNs). It is represented as an extension of CNN, with an additional transformational quanvolutional layer. Following are the four steps to transform classical data into different feature maps using quanvolutional filters:

- The quanvolutional filter selects a small subsections (u_x) of an input image (u) and applies a random quantum circuit (Q), where each u_x is a 2-d matrix.
- There are different ways to encode the small subsections (u_x) into a quantum state (q). We selected encoding function (E) and initialization of quantum state is defined as $q_x = E(u_x)$.
- After encoding into the quantum state, the next step is to apply the quantum circuit (Q). The outcome is defined as $f_x = Q(q_x) = Q(E(u_x))$.
- Last step is to decode the information by applying the finite number of measurements in the form of a decoding function (D). The final outcome (o_x) decoded state is defined as $o_x = D(f_x) = D(Q(E(u_x)))$, where o_x represents a scalar value. The transformation of classical data into different feature maps using quanvolutional filter transformation (Q_f) is represented as $o_x = Q_f(E, Q, D)$. Repeating the same steps over different subsections to traverse the full input image, and produce an outcome that will be organized as a multi-channel image.

The segment of an input image is encoded into a quantum state by the quanvolutional filters containing a random parameterized quantum circuit, as shown in figure 1. Instead of applying element-wise matrix multiplications, a small subsection (2×2) of an input image corresponding to filter size is encoded into a quantum state $|q_x\rangle$ by applying a series of parameterized rotational one-qubit gates such as $R_i(\theta)$, where

Figure 1. Example of a quantum convolutional layer in quanvolutional neural network. The 2×2 part of an image is encoded into a variational quantum circuit by giving parameterized rotations.

$i \in \{X, Y, Z\}$. Due to the limited width of the quantum hardware, the quanvolutional operation on the whole image at once is unfeasible. A hybrid QCNNs are designed for the NISQ era. For simplicity, we used random quantum circuits for quanvolutional layers with parameters $\theta = (\theta_1, \theta_2, \ldots, \theta_j)$ to produce feature maps as an output. A final measurement is performed on a computational basis, estimating four expectation values, which are mapped into four different channels. The resulting features from the quantum circuit are fed into the classical layer preceding with pooling layers. Then, the subsequent measurement is performed for detecting meaningful features of the segments, which are processed by a fully-connected classical layer (F) with parameters $p = (p_1, p_2, \ldots, p_k)$ for the classification task [24]. Therefore, the complete process after applying the classical fully connected layer is defined as $o_x = Q_f(E, Q, D, F)$. Using quanvolutional layers, quantum computers can analyze the data in high-dimensional Hilbert spaces more accurately than classical computers. It is very crucial to design a parameterized quantum circuit in which the number of qubits is identical to the filter size. Motivated by the architectural design of QCNN, we extend its potential to quantum federated learning (QFL) with collaborative training on QCNNs.

With the rapidly growing concerns of the above-mentioned two approaches, it is a natural objective to unlock the potential to take artificial intelligence to its next generation without centralized training data. Healthcare data is a valuable asset. Due to highly sensitive medical data, it is crucial to deploy distributed machine learning algorithms for enabling scalable and privacy-preserving intelligent healthcare applications.

Coronavirus Disease 2019 (COVID-19) is still a serious life-threatening disease and federated learning is expected to become an effective tool for detecting COVID-19 without sharing patient's data between hospitals. Data privacy is a major concern and several organizations are reluctant to share data or participate in big data communities, which prevents them to gain an advantage from data analytical methodologies. Academic researchers and industry practitioners are unaware of the value it could offer them. Therefore, a mechanism needs to be designed to tackle the optimization, security, and privacy challenges of highly sensitive medical data. This paper provided the QFL framework and its potential for efficient training on publicly available medical datasets COVID-19 [25] and MedNIST [26], which are distributed among several parties. The purpose is to facilitate confidence in different sectors for information sharing using quantum technological solutions. To the best of our knowledge, no hybrid quantum–classical or classical-quantum

CNNs have been proposed that attempt to implement federated learning over non-independent and non-identical distribution (Non-IID) and real-world distribution of data among quantum nodes.

1.1. Aim and contributions
To summarize, this paper makes the following contributions:

- Proposed a QFL framework based on a QCNN that works collaboratively with several quantum nodes without exchanging the data.
- Investigated the performance of proposed QFL framework on medical classical machine learning datasets, COVID-19 and MedNIST, which are distributed non-independent and non-identically among the quantum nodes and real-world distribution of datasets.
- The effectiveness of federated QCNN (FedQCNN) and federated classical-quantum CNN (FedCQNN) is analyzed when the clients have insufficient training samples.
- When clients have insufficient training samples, the federated models require more rounds to achieve the target accuracy in real-world distribution compared to the Non-IID partitioned MedNIST dataset.

The remainder of this paper is organized as follows: section 2 is the related work. Section 3 contains the description of QFL and the concept of federated QCNNs. In section 4, the problem formulation with medical dataset description is given. Section 5 presents the experimental results and discussions on the findings and outcomes. Finally, the concluding remarks are reported in section 6.

2. Related work

A handful of related works exist on collaborative quantum machine learning models. The concept of QFL was first studied by Chen and Yoo [27] with an objective to enhance data privacy using a hybrid quantum–classical machine learning model in distributed settings. The collaborative training is performed on hybrid quantum–classical classifiers and investigated that it converges in fewer communication rounds as compared to centralized models. Li *et al* [28] proposed a private quantum protocol for single and multi-party distributed learning using blind quantum computing. It has been demonstrated that the proposed protocol incorporated with differential privacy has the potential to handle large-scale distributed learning tasks with security guarantees.

Xia and Li [29] proposed the concept of a QFL framework (QuantumFed) with collaborative training on quantum neural networks. Later, Xia *et al* focused on Byzantine-resilient algorithms in quantum federated and compared the performance with classical distributed learning systems. Yamany and [30] deployed the QFL framework in intelligent transportation systems. The quantum-behaved particle swarm optimization algorithm is used to update the hyperparameters and has proven to be resilient against adversarial attacks in federated settings.

Recently, Chehimi and Saad [31] introduced the QFL framework to operate on quantum data in quantum networks. The quantum version of the federated dataset is generated and performed a binary classification using eight qubits quantum CNNs. It has been demonstrated that the proposed framework achieved efficient training on independent and identically distributed (IID) and Non-IID distributed datasets. Huang *et al* [32] proposed the communication efficient QFL framework based on variational quantum tensor networks, quantum approximate optimization algorithm, and variational quantum eigensolver. It has been observed that the proposed framework reduced the possible attacks on client data and gained an exceptional performance in variational tasks. Zhang *et al* [33] proposed a quantum entangled state-based secure aggregation (QSA) scheme for securing private quantum model parameters against eavesdropping. The proposed scheme has shown versatility by implementing logistic regression, CNNs, and quantum neural networks. Recently, Bhatia *et al* [34, 35] introduced a QFL framework consisting of quantum CNNs for the healthcare sector. The authors showcased the effectiveness of the proposed framework across various medical data distributions and asserted its ability to achieve comparable performance with established classical models for pneumonia and CT-kidney datasets. Moreover, a communication-efficient client selection strategy is proposed to reduce the rounds and to accelerate convergence within the QFL framework.

3. QFL optimization

In this section, we will focus on how healthcare organizations can leverage the potential of QFL. Healthcare organizations/hospitals are trying to solve how large highly-sensitive data can be used to enhance their services. QFL aims to take the benefit of multiple clients that can commit to global tasks independently.
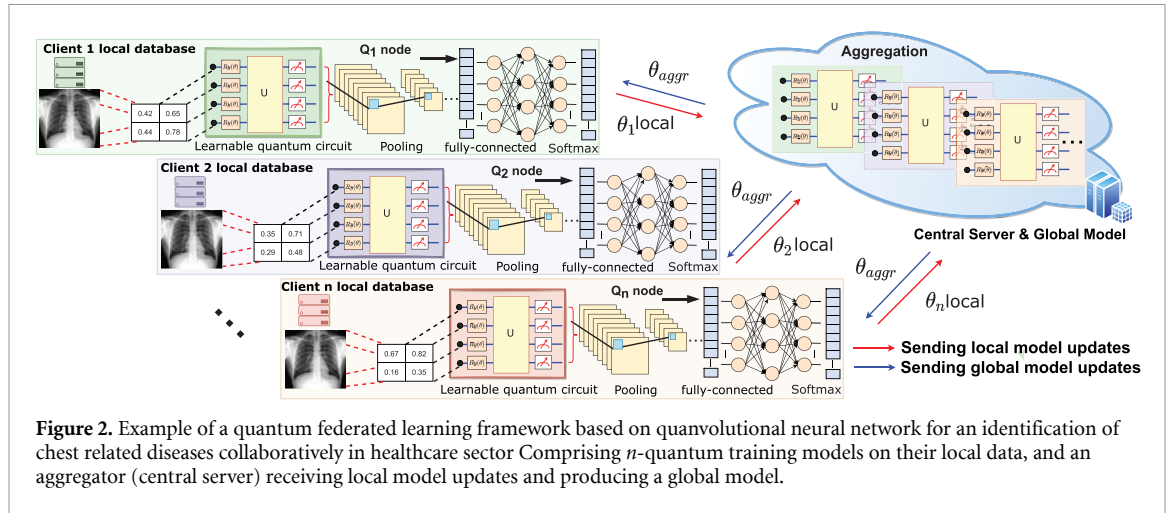
Figure 2. Example of a quantum federated learning framework based on quanvolutional neural network for an identification of chest related diseases collaboratively in healthcare sector Comprising *n*-quantum training models on their local data, and an aggregator (central server) receiving local model updates and producing a global model.

Motivated by the potential of federated learning and quantum machine learning, we proposed the concept of a federated QCNN framework with decentralized data, as shown in figure 2. It consists of *K*-local clients (or healthcare organizations/hospitals) and train a quantum machine learning model (QCNN) on their local medical datasets. All local clients train their local models within each communication round. Then, the healthcare organizations send their model parameters, while the local data stays on the client side that owns it. Finally, the aggregation is performed of all local model updates and a global server sends the updated model weights to all organizations for further training.

Even relatively small-scale healthcare agencies can participate and take benefit from the collaborative models generated. It allows individual hospitals, research institutions, and agencies to benefit from the rich datasets without centralizing the data in one place. It has the potential to mitigate major problems such as data security, and data access rights. The client's privacy protection can be improved by performing the training of models where the local data sources are, instead of sending raw datasets to the centralized server.

3.1. FedQCNN
3.1.1. Local QCNN models
We considered a FedQCNN setup consisting of a global model and a set of clients sharing a QCNN. Assume there is a set of *K* quantum computing nodes/clients/clinical institution data warehouse where the data reside in. Let n_k denote the number of samples available from client (k) and $N = \sum_{k=1}^{K} n_k$ is the total sample size, where $k \in K$. The complete steps of federated quantum learning averaging algorithm are represented by pseudo-code in algorithm.

At each client side, the entries of unitary matrices are the learnable parameters (θ) generated by the quantum circuit. Let $\theta^k = (u_x, E, Q, D, F))$ denotes the vector of all quanvolutional parameters and QCNN transforms the initialized zero-state quantum registers into image-state $|q\rangle$ by selecting a small subsection u_x. The predicted outcome by QCNN is represented by $f_{\theta_k}|q\rangle$ for client k, where q image of size $r \times s$ pixels after applying encoding function (E) consisting of parameterized rotations. For encoding, only quadratic subsections of an input image are selected. The quantum circuit parameters are updated to minimize the mean squared error.

$$\arg\min F\left(\theta^k\right) = \frac{1}{n_k} \sum_{i=0}^{n_k} \left(y_i - \left(f_{\theta^k}|q_i\rangle\right)\right)^2 \qquad (1)$$

where n_k is the number of samples associated with client *k*. In the FedQCNN setup, collaborative training is achieved by sending the classical parameters from local clients to the central server. In our experiments, all *K* quantum nodes/clients participated in collaborative training of the global QCNN model.

Algorithm 1. Federated learning averaging algorithm.

---

1. **Input:** K clients such that $k = 1, 2, \ldots, K$, E defines the number of local epochs, B represents batch-size, T states communication rounds, n_k is the number of samples available with client k and η_s, η_c are the learning rates of server and client, respectively.
2. **Server Aggregation:**
3. Initialize $\theta_{1\text{shared}}$
4. **for** $t = 0$ to $T\text{-}1$ **do**
5.   Send θ_t global server parameters to each client.
6.   **for** each $k \in K$ **do**
7. $\theta^k_{t+1\text{local}} \leftarrow \text{Clientupdate}(k, \theta^k_{t\text{shared}})$,
8.   **end for**
9. $\theta_{t+1\text{shared}} \leftarrow \theta^k_{t\text{local}} - \eta_s \frac{1}{K} \sum_{k=1}^{K} \theta^k_{t+1\text{local}}$
10. **end for**
11. **Clientupdate**(k, θ):
12. **for** $i = 1$ to E **do**
13.   **for** $b \in B$ **do**
14. $\theta^k \leftarrow \theta^k - \eta_c \nabla L_k(\theta^k, b)$
15.   **end for**
16. **end for**
17. return θ to Server

3.1.2. Global model

At the beginning of QFL, the central server holds a global QCNN model and the quantum nodes or local clients hold their local datasets. In each communication round (t), the central server sends the global model QCNN parameters (θ_t^s) to all K clients for a collaborative training. Each client k trains the current global QCNN model's parameters on their local data and executes a local stochastic gradient descent (SGD) for optimization with a fixed learning rate η_c to generate the updated parameters as

$$g^k = \nabla_{\theta_t^k} F\left(\theta_t^k\right) \tag{2}$$

$$\theta_t^k = \theta_t^s - \eta_c g^k. \tag{3}$$

Once the client (k) finishes the current round (t) of training, the updated parameters are uploaded back to the server for aggregation. Finally, the central server then averages the local model parameters to update its global model by employing a federated averaging algorithm and sends the new updated set of parameters $\theta^k_{t+1\text{shared}}$ to all k clients for the next communication round ($t+1$).

$$\theta_{t+1\text{shared}} \leftarrow \theta^k_{t\text{local}} - \eta_s \frac{1}{K} \sum_{k=1}^{K} \theta^k_{t+1\text{local}} \tag{4}$$

where θ^k_{t+1} denotes the weight vector at communication round ($t+1$) and η_s is the learning rate of server. This parameters exchange process repeats between a central server and local clients until it converges after a certain amount of communication rounds. The result is an optimal QCNN model that can be customized for personalization.

3.1.3. Federated classical-quantum CNN (FedCQNN)

Next, we considered a federated version of a hybrid classical-quantum CNN (FedCQNN). In our experiments, the classical layer consists of one CNN layer followed by a fully-connected dense layer. The outcome from the classical CNN is collected and embedded into a quantum state at the input interface of the quantum layer. The quantum layer is built with a variational quantum circuit (VQC) to create a hybrid CNN for the classification task, as shown in figure 3(A). The features are extracted from the classical layer and are encoded into 16-qubits using angle encoding in the quantum layer. The qubits at features with values that exceed a threshold (i.e. 0.5) are rotated about the X-gate and Y-gate.

The classical output vector is transformed into a 16-qubit quantum state ($q_0, q_1, \ldots, q_{15}$), which is given as an input to the VQC, and one qubit is used as a readout qubit for binary classification, i.e. labeled as (-1, -1). It is prepared using one-qubit X and Hadamard (H) gates. The classification outcome is based on the expectation value of the readout qubit in a parameterized quantum circuit. It returns a value between -1 and 1. The VQC is built using a sequence of XX (i.e. tensor product of two X gates) and ZZ gates (i.e. tensor
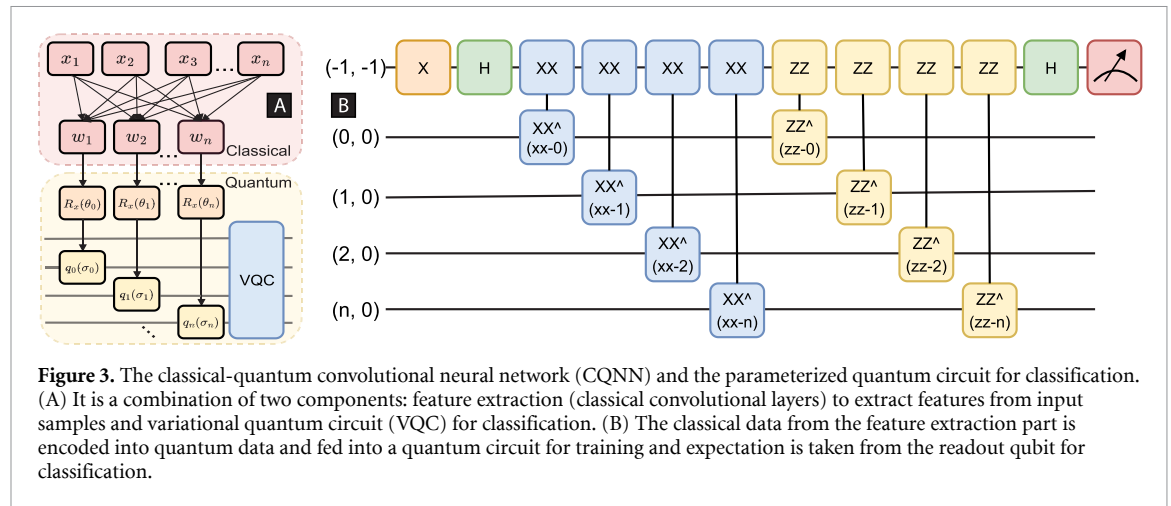
Figure 3. The classical-quantum convolutional neural network (CQNN) and the parameterized quantum circuit for classification. (A) It is a combination of two components: feature extraction (classical convolutional layers) to extract features from input samples and variational quantum circuit (VQC) for classification. (B) The classical data from the feature extraction part is encoded into quantum data and fed into a quantum circuit for training and expectation is taken from the readout qubit for classification.

product of two Z gates), with all the data qubits (i.e. labeled as (0, 0), (1, 0), …, (15, 0)) acting on the readout qubit, as shown in figure 3(B).

In the FedCQNN setup, the collaborative training is achieved by sending the classical parameters from local CQNN models to the central server holding a global CQNN model.

4. Problem formulation

Here, we investigated the feasibility of federated QCNNs and federated classical-quantum CNNs with Covid-19 and MedNIST datasets. Initially, we focused on non-independent and non-identically distributed (Non-IID) partitioning of data, where the data is shuffled and distributed to clients. Further, we consider the real-world distribution of data among the clients to replicate the real-world scenario.

4.1. Dataset description
For QFL benchmarks, we collected the publicly available CC-19 dataset. It contains 34 006 computed tomography (CT) scan images of 89 subjects, which consists of 5611 ct-scan slices of negative COVID-19 patients and 28 395 positive COVID-19 patients. For the multi-classification task in radiology and healthcare imaging settings, the additional experiments are performed on the publicly available MedNIST dataset. It consists of 58 954 medical images, created by collecting different medical imaging datasets from TCIA, the NIH Chest x-ray dataset, and the RSNA bone age challenge. MedNIST dataset contains six classes of biomedical images and each consisting of 10 000 images: Abdomen CT, Breast MRI, Chest CT, Chest XR, Hand, and Head CT. Some random samples of Covid-19 and MedNIST datasets are shown in figure 4. All the experiments are performed using a quantum programming framework Tensorflow Quantum (TFQ) by Google [36, 37].

5. Experiments and discussion

To demonstrate the effectiveness of federated QCNN and CQNN, we studied two Non-IID data partitioning strategies in the experiments, named Non-IID1 and Non-IID2. All the processing algorithms and the architecture of federated QCNN and CQNN are kept equal to support a comparable pipeline. Firstly, the dataset is randomly distributed to the given number of clients. The number of samples of each class may vary for each client. Therefore, the mean of the whole sample size of each client is different from that of other clients. In Non-IID2, each client can have random samples of at least three classes of a dataset. The comparison of federated model architectures between QCNN and CQNN is given in a table 1. To study the federated optimization in quantum settings, we experiment with federated averaging (FedAVG) and federated SGD (FedSGD). In FedSGD, the batch size (B = ∞) and the client (C = 1) can perform one local SGD update per round of communication.

Our initial study includes, the Non-IID1 distribution of COVID-19 dataset for binary classification and MedNIST dataset for multi-classification tasks. The federated QCNN consists of a single quanvolutional layer of depth one for prepossessing of a dataset, followed by a pooling operation and a fully-connected layer with a final softmax activation function. The federated CQNN consists of single convolutional layer, followed by a pooling operation. Afterward, the resulting features are fed into a 16-qubit VQC for classification.
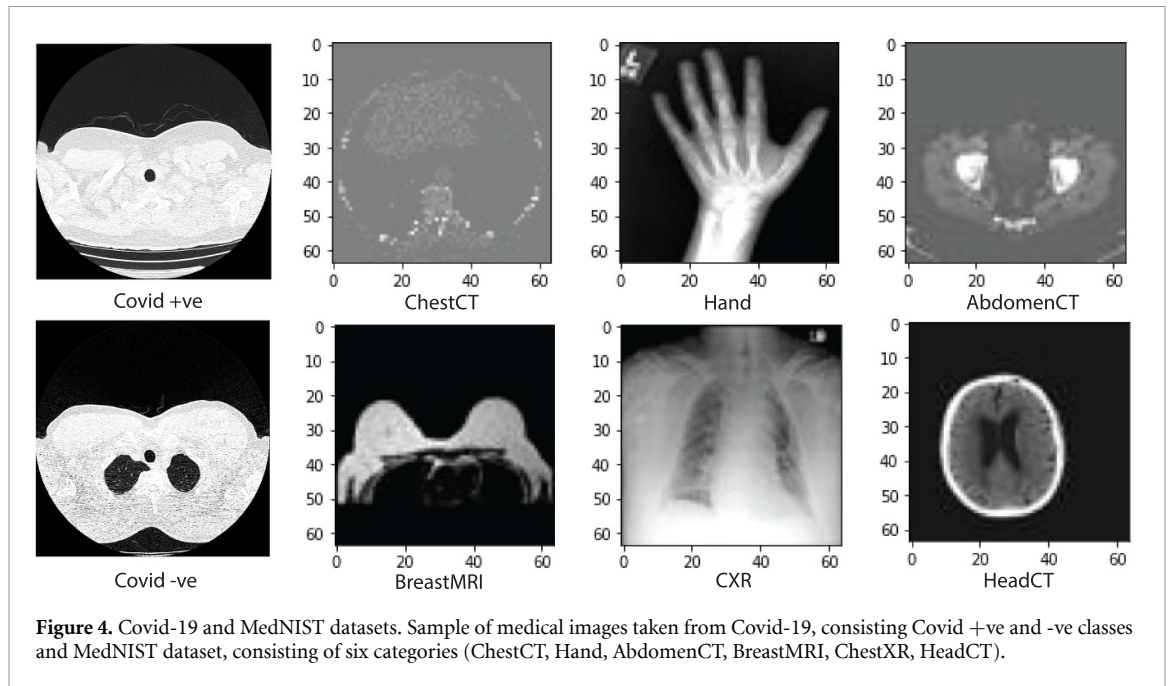
Figure 4. Covid-19 and MedNIST datasets. Sample of medical images taken from Covid-19, consisting Covid +ve and -ve classes and MedNIST dataset, consisting of six categories (ChestCT, Hand, AbdomenCT, BreastMRI, ChestXR, HeadCT).

Table 1. Comparison of federated model architectures between QCNN and CQNN.

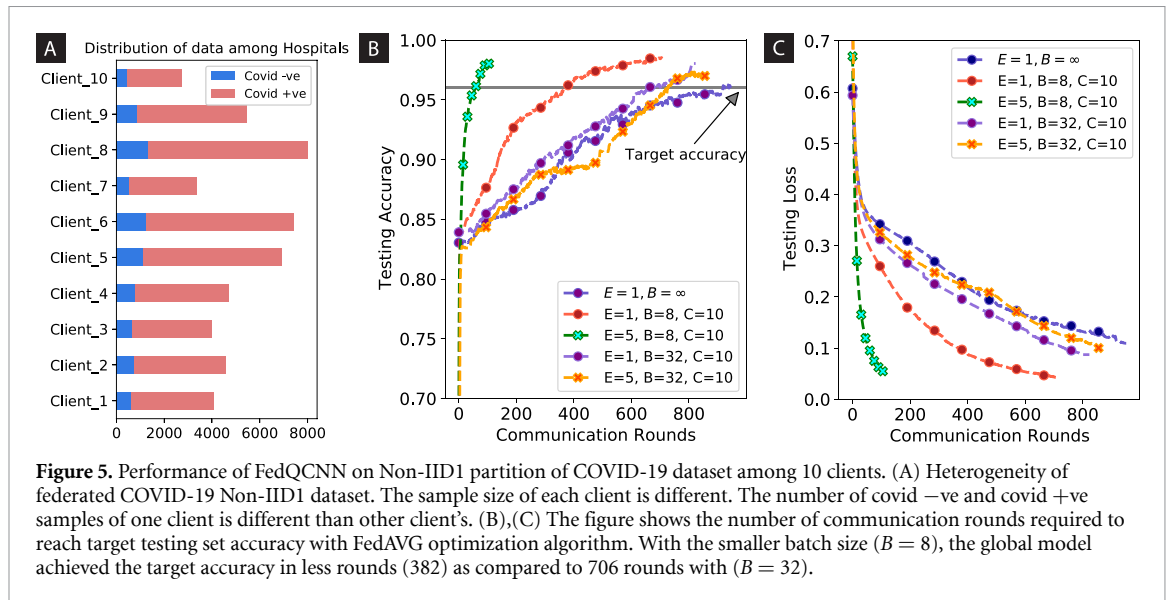| Federated hybrid models	No. of channels/Filter size/FC layer	Execution time (in sec) per epoch (10 clients)	Total number of parameters
FedCQNN	$4/2\times2/16$	2.041	10 962
FedQCNN	$1/2\times2/16$	6.027	11 788

Figure 5. Performance of FedQCNN on Non-IID1 partition of COVID-19 dataset among 10 clients. (A) Heterogeneity of federated COVID-19 Non-IID1 dataset. The sample size of each client is different. The number of covid −ve and covid +ve samples of one client is different than other client's. (B),(C) The figure shows the number of communication rounds required to reach target testing set accuracy with FedAVG optimization algorithm. With the smaller batch size ($B = 8$), the global model achieved the target accuracy in less rounds (382) as compared to 706 rounds with ($B = 32$).

In the first experiment, we extend our QFL framework to binary-class classification task. The Non-IID1 distribution of the COVID-19 dataset among ($C = 10$) clients is shown in figure 5(A). The testing set accuracy 96% is achieved by executing QCNN in a non-federated fashion. We address the communication rounds required to reach the target testing set accuracy with a federated QCNN. For all the experiments on COVID-19 dataset, the client and server learning rates are set as 0.01 and 0.1, respectively. Figures 5(B) and (C) shows the testing accuracy and loss versus the number of communication rounds. The gray line indicates the target testing set accuracy. The federated QCNN with FedSGD algorithm ($B = \infty$) reaches the 96% accuracy after 951 communication rounds. Table 2 entries give the number of communication rounds needed to reach target testing set accuracy, besides the speedup respective to FedSGD optimization

Table 2. Number of communication rounds taken by FedQCNN to reach the target accuracy. For Non-IID1 partition of COVID-19 dataset, adding more computation on each local quantum node decreases the rounds of communication to achieve a testing set accuracy.

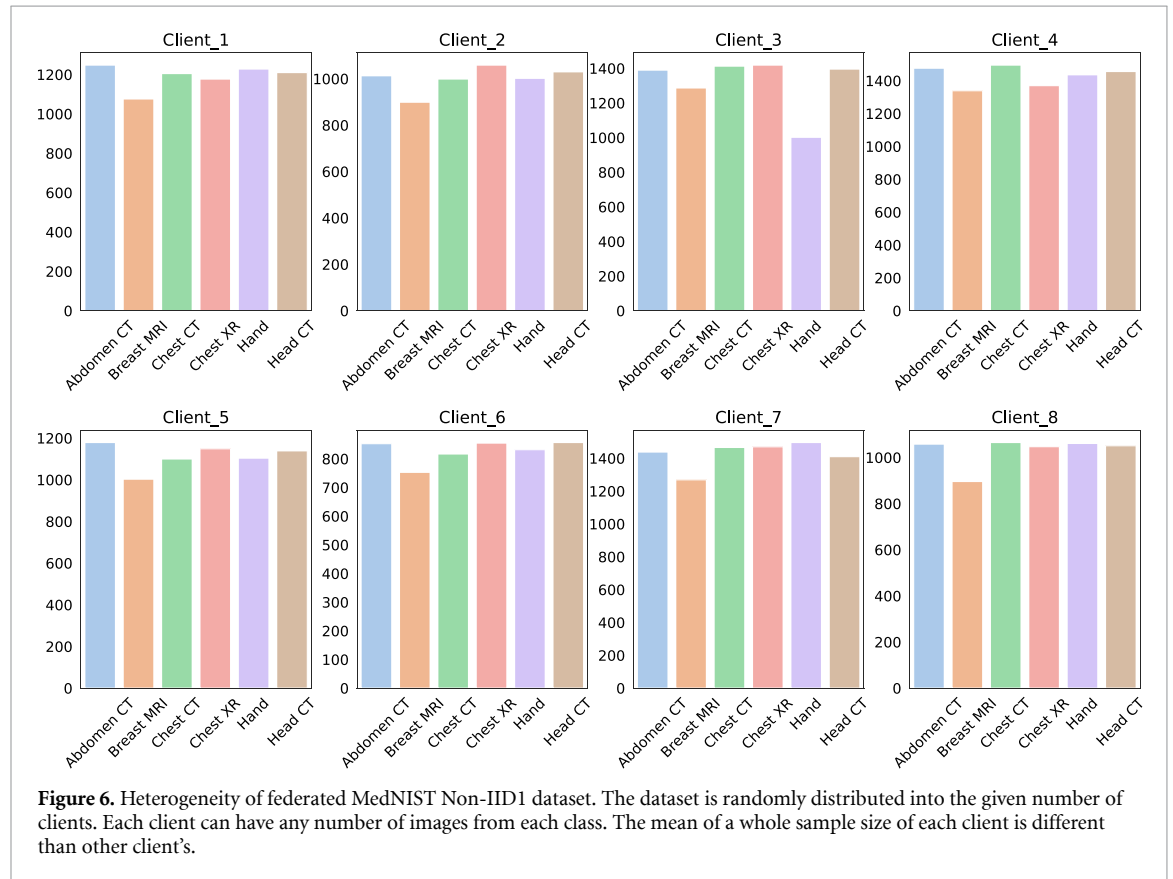|	COVID-19 Non-IID1 dataset (96%)			
Fed	C	E	B	QCNN
FedSGD		1	∞	951
FedAVG	10	1	8	382 (2.44$\times$)
FedAVG	10	5	8	65 (14.12$\times$)
FedAVG	10	1	32	706 (1.34$\times$)
FedAVG	10	5	32	731 (1.31$\times$)

Figure 6. Heterogeneity of federated MedNIST Non-IID1 dataset. The dataset is randomly distributed into the given number of clients. Each client can have any number of images from each class. The mean of a whole sample size of each client is different than other client's.

algorithm. On increasing the number of local epochs (E), it has produced a significant reduction in communication rounds with (B = 8).

We then extend the QFL framework to multi-class classification tasks to evaluate its robustness in distributing the MedNIST dataset randomly among 10, 20, and 50 clients. An example of its Non-IID1 distribution among the first eight clients is shown in figure 6, where each client can have any number of images from each class. For all the experiments on the MedNIST dataset, the learning rate for client and server is set as 0.02. The speedups in communication rounds are corresponding to the FedQCNN with FedSGD optimization. Figure 7 shows the impact of varying number of local epochs (E) and batch-size (B) with clients. The number of communication rounds is calculated when the curve crosses the target testing set 97% accuracy, which is represented by a gray line.

The federated QCNN with FedSGD algorithm (blue color curve) eventually reaches the target testing set accuracy after 1950 communication rounds. Table 3 shows the effect of client fraction on the MedNIST dataset with a federated QCNN. There seems to be some advantage in increasing the number of clients. In the Non-IID1 case of MedNIST, the federated QCNN with 20 clients has reduced the communication rounds by 6 times as compared to 10 clients. The quantum federated model with FedAVG algorithm converges to a higher testing accuracy in less communication rounds as compared to FedSGD model. On increasing the fraction of clients, the federated global model still robustly achieves a target set accuracy. With a smaller batch size, FedQCNN shows a good stability between convergence rate and computational efficiency.
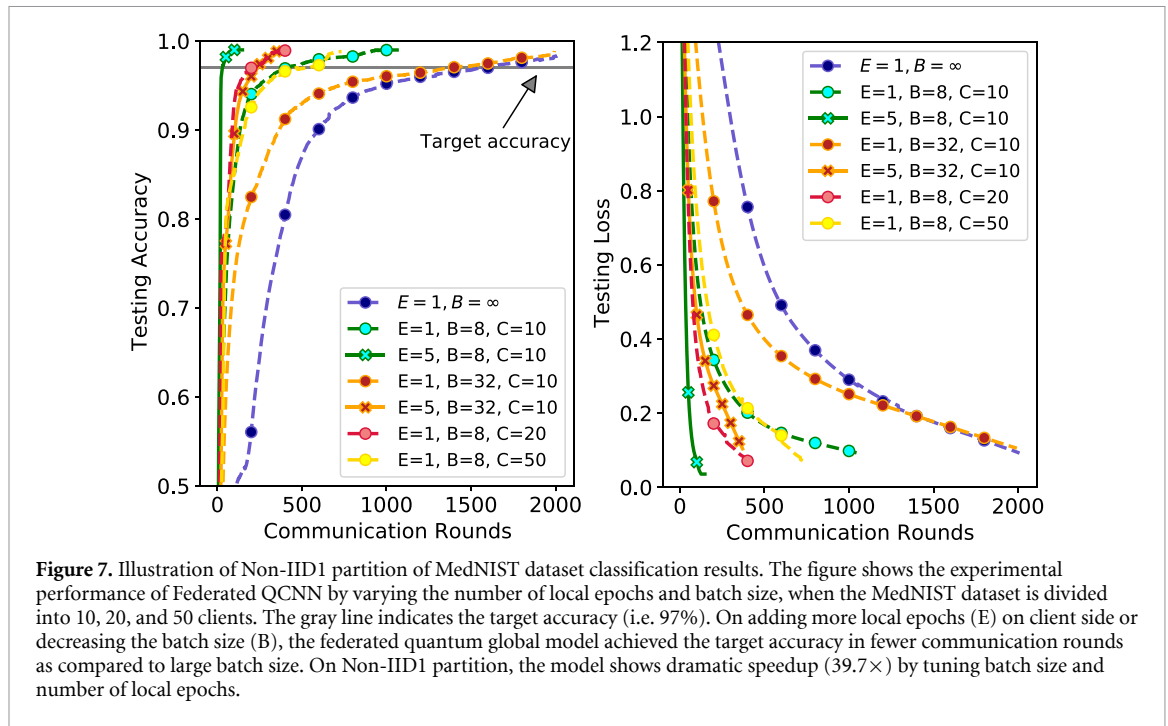
Figure 7. Illustration of Non-IID1 partition of MedNIST dataset classification results. The figure shows the experimental performance of Federated QCNN by varying the number of local epochs and batch size, when the MedNIST dataset is divided into 10, 20, and 50 clients. The gray line indicates the target accuracy (i.e. 97%). On adding more local epochs (E) on client side or decreasing the batch size (B), the federated quantum global model achieved the target accuracy in fewer communication rounds as compared to large batch size. On Non-IID1 partition, the model shows dramatic speedup (39.7×) by tuning batch size and number of local epochs.

Table 3. Effect of client (C) division on Non-IID1 distribution of MedNIST dataset. The global FedQCNN model with 20 clients has reduced the communication rounds by six times as compared to 10 clients.

	MedNIST Non-IID1 dataset (97%)			
Fed	C	E	B	QCNN
FedSGD		1	∞	1950
FedAVG	10	1	8	626 (3.11×)
FedAVG	10	5	8	49 (39.7×)
FedAVG	10	1	32	1750 (1.11×)
FedAVG	10	5	32	289 (6.74×)
FedAVG	20	1	8	286 (6.81×)
FedAVG	50	1	8	646 (3.01×)

We then turn to examine the robustness and stability of FedQCNN with a significantly insufficient and uneven dataset distributions. We considered the Non-IID2 partition of MedNIST, where each client can have random images from at least three classes. We investigated the performance of FedQCNN and FedCQNN with FedAVG optimization algorithm. The distribution of MedNIST dataset among 10 clients is shown in figure 8(A). Firstly, from table 4, we can observe that the federated hybrid algorithms did not reach the target testing set accuracy in the allowed time with $B = 32$. Second, in Non-IID2 settings, when number of clients possess insufficient training samples, the federated models require a higher number of rounds to reach target set accuracy, as shown in figures 8(B) and (C). Third, the effect of training for several local epochs (E) between federated averaging can be seen. On dividing the dataset among 20 clients and fixing $E = 5, B = 8$, the FedQCNN reduces the communication rounds by 4.9 times, as compared to 1.3 times with $E = 1$. Figure figures 8(D) and (E) shows the accuracy of FedQCNN with the confusion matrix after 100 communication rounds with $E = 1, 5$. It has been observed that 20 clients progresses their learning slowly with a single local SGD update. The federated QCNN is confused in making predictions between class Abdomen CT and Chest CT. The generated confusion matrix with $E = 5$ correctly classified the relevant classes after 100 rounds. Even with $B = 32$ and $E = 5$, the federated QCNN and CQNN decrease the communication rounds dramatically to reach target set accuracy.

Finally, we performed the experiments on real-world federated dataset to validate the results of federated QCNN and CQNN. The real-world distribution among 10 clients is shown in figure 9(A), where each client can have random images from any number of classes of MedNIST dataset. We have the following observations in aspect of different experiments. First, from table 5, we observed that the federated hybrid models not able to achieve target testing accuracy with 10 clients in the allowed time based on real-world distribution. As depicted in figure 9(A), each client has notably inadequate number of training samples, with
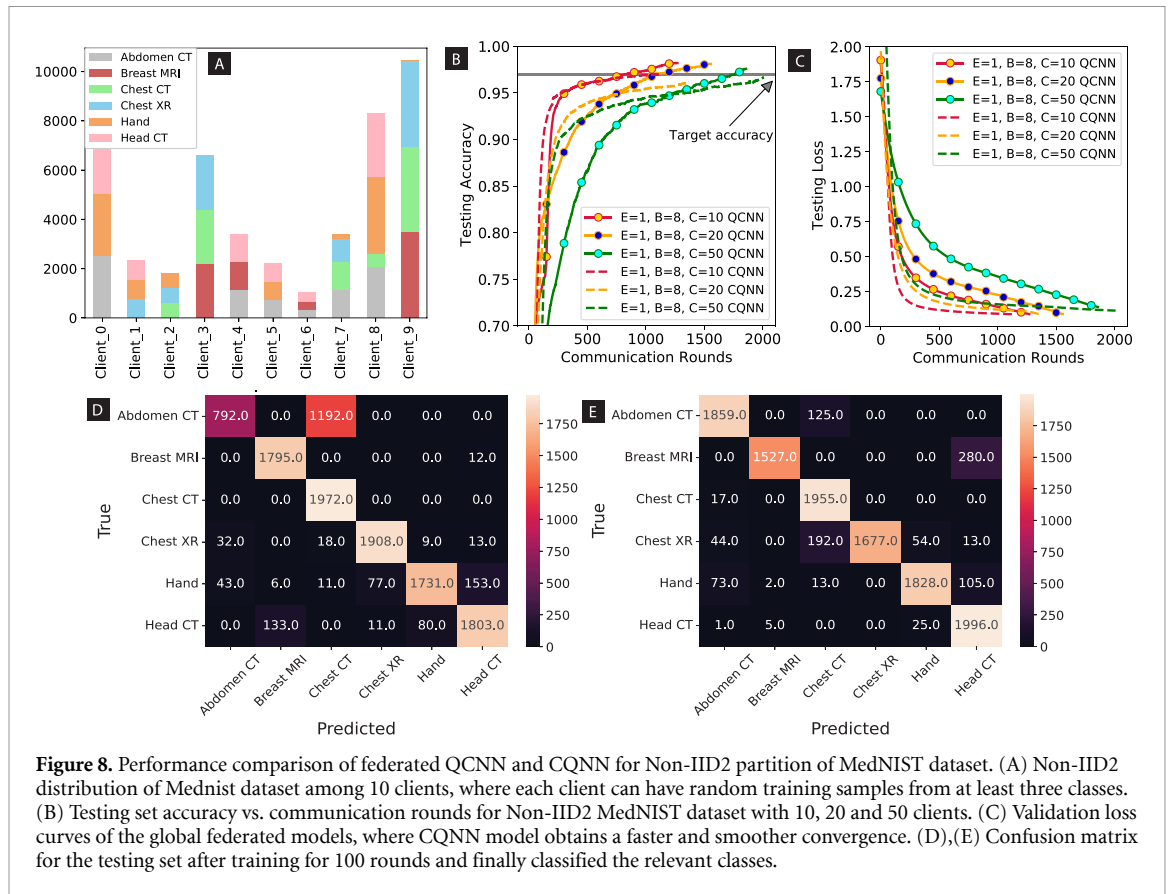
Figure 8. Performance comparison of federated QCNN and CQNN for Non-IID2 partition of MedNIST dataset. (A) Non-IID2 distribution of Mednist dataset among 10 clients, where each client can have random training samples from at least three classes. (B) Testing set accuracy vs. communication rounds for Non-IID2 MedNIST dataset with 10, 20 and 50 clients. (C) Validation loss curves of the global federated models, where CQNN model obtains a faster and smoother convergence. (D),(E) Confusion matrix for the testing set after training for 100 rounds and finally classified the relevant classes.

Table 4. Number of communication rounds by FedQCNN and FedCQNN versus testing set accuracy. When the clients have insufficient training samples in Non-IID2 distribution of MedNIST, the federated models still show substantial speedup on achieving the target set accuracy.

MedNIST dataset				
			Non-IID2 distribution	
C	E	B	QCNN	CQNN
10	1	8	1152 (1.69×)	1139 (1.72×)
20	1	8	1493 (1.30×)	1583 (1.23×)
50	1	8	1896 (1.02×)	>2000 (0.74×)
20	5	8	397 (4.91×)	927 (2.10×)
10	1	32	>2000 (0.64×)	>2000 (0.47×)
10	5	32	624 (3.12×)	959 (2.03×)

some even (client1 and client7) possessing training samples from only one class. Second, the FedCQNN achieves best accuracy in less number of communication rounds with 20 clients and $E = 5$, as shown in figures 9(B) and (C). Third, if B is large enough, the federated hybrid models not able to reduce the computational cost. Five runs with $B = 32$, the FedQCNN and FedCQNN decrease the communication rounds by 2.8 and 2.0 times, respectively and reach the target testing set accuracy. The confusion matrix of FedQCNN is visualized for 20 clients fixing $B = 8$ and $E = 1, 5$ in figures 9(D) and (E). After 100 rounds, the FedQCNN ($E = 1, B = 8, C = 20$) not able to classify certain classes (chest CT and Abdomen CT) correctly. On increasing the local SGD updates, the counts from predicted and actual values of chest CT class has been increased significantly after adding more communication rounds, as shown in the confusion matrix figure 9(E).

5.1. Adding computation per client

An important aspect of active research is the reduction of communication burden throughout the federated learning process. This plays a vital role in enhancing effectiveness and curtailing the total training duration. Several strategies can be utilized to accomplish this objective, including increasing local training epochs,
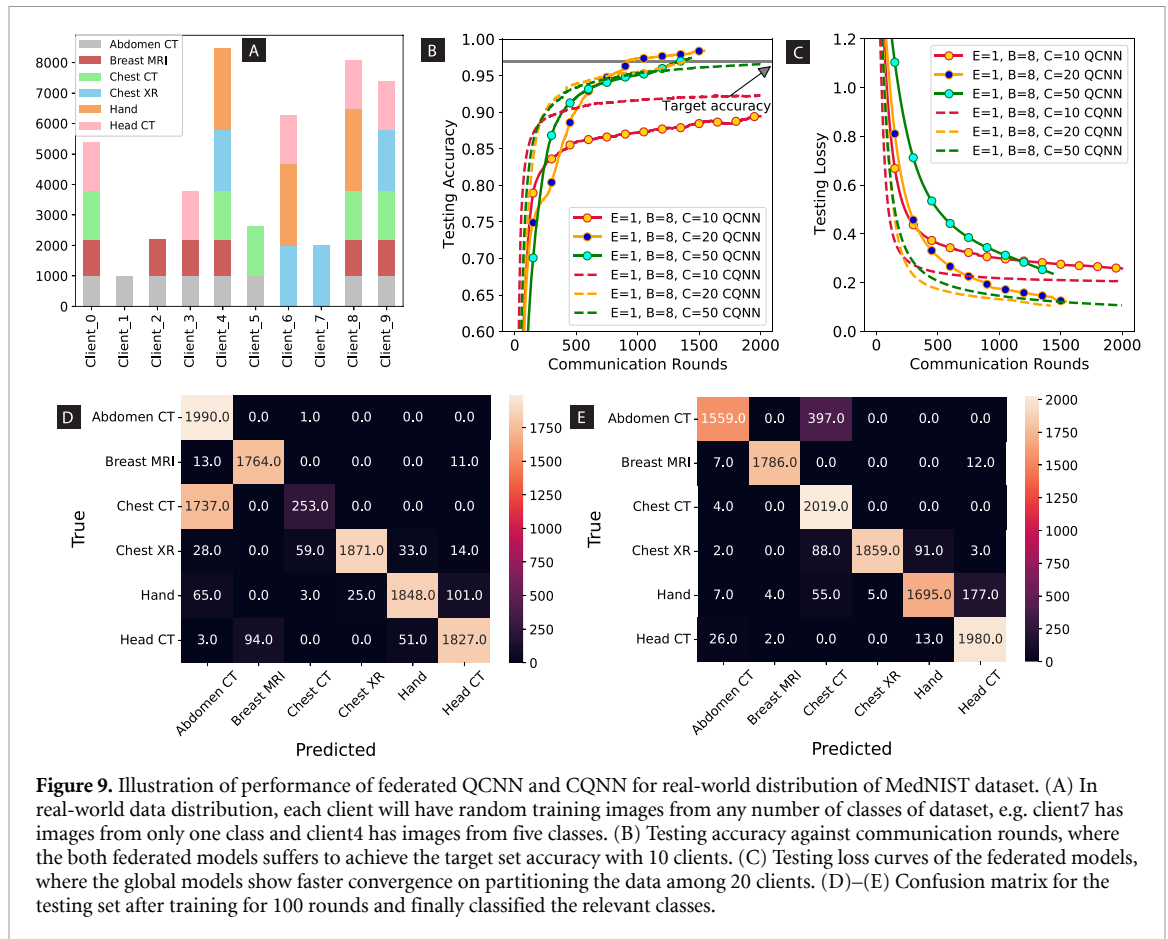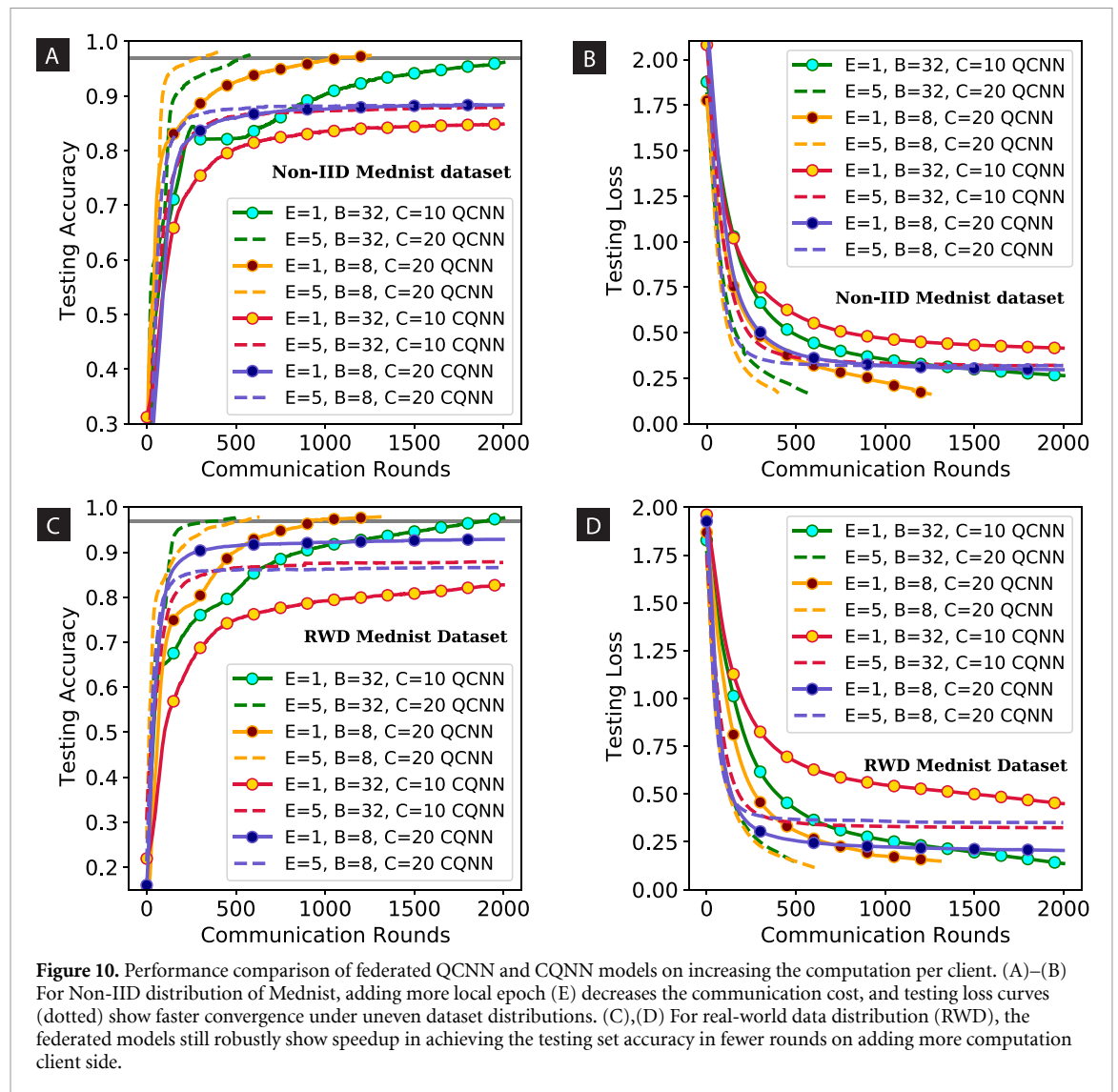
Figure 9. Illustration of performance of federated QCNN and CQNN for real-world distribution of MedNIST dataset. (A) In real-world data distribution, each client will have random training images from any number of classes of dataset, e.g. client7 has images from only one class and client4 has images from five classes. (B) Testing accuracy against communication rounds, where the both federated models suffers to achieve the target set accuracy with 10 clients. (C) Testing loss curves of the federated models, where the global models show faster convergence on partitioning the data among 20 clients. (D)–(E) Confusion matrix for the testing set after training for 100 rounds and finally classified the relevant classes.

Table 5. Performance comparison of FedQCNN and FedCQNN for real-world distribution of MedNIST dataset. Both models show substantial smaller speedups in achieving the testing set accuracy.

MedNIST dataset				
			Real-world distribution	
C	E	B	QCNN	CQNN
10	1	8	>2000 (0.27×)	>2000 (0.37×)
20	1	8	1543 (1.26×)	1421 (1.37×)
50	1	8	>2000 (0.97×)	>2000 (0.72×)
20	5	8	774 (2.51×)	489 (3.98×)
10	1	32	>2000 (0.88×)	>2000 (0.36×)
10	5	32	685 (2.84×)	966 (2.01×)

selectively involving clients, adaptive learning rates, optimizing batch sizes, implementing model compression, and many more.

Our emphasis centered on two primary tactics, we are adding up the computation on the client side by either increasing local epochs, decreasing batch size, or both to reduce the communication rounds in a QFL framework. The distribution of the data is based on Non-IID2 and real-world settings among 10 and 20 clients. The solid curves show the results with a single local SGD update and dotted curves represent the experiments with $E = 5$. The results of both partitioning strategies by varying B and E are shown in figures 10(A)–(D). It has been investigated that increasing the number of local epochs on the client side can reduce the communication rounds on the testing set significantly. The speedups for the real-world partitioned dataset are smaller as compared to Non-IID2 with FedQCNN, but still considerable $(2.5 - 2.9\times)$, as shown in figures 10(C) and (D). In fact, it can be due to some clients having large local datasets in Non-IID2, which makes a valuable increment in local training. Although, the fewer local epochs per round results in slow convergence. On increasing B, the federated hybrid models are not able to achieve the target testing set accuracy. Adding the number of local epochs on each training quantum node with a larger B helps to lower the communication rounds. Therefore, the tuning of a batch-size parameter is

Figure 10. Performance comparison of federated QCNN and CQNN models on increasing the computation per client. (A)–(B) For Non-IID distribution of Mednist, adding more local epoch (E) decreases the communication cost, and testing loss curves (dotted) show faster convergence under uneven dataset distributions. (C),(D) For real-world data distribution (RWD), the federated models still robustly show speedup in achieving the testing set accuracy in fewer rounds on adding more computation client side.

essential to maintain a balance between computational efficiency and convergence rate. Moreover, the decrease in learning rate with the increase in local epochs can be crucial in the final stages of training.

6. Conclusions and future work

In this work, we proposed a novel framework to train QCNNs in federated learning settings that allows the implementation of collaborative quantum learning, where the data is distributed randomly among the clients. We assessed the reliability and performance of federated QCNNs and federated classical-quantum CNNs on medical datasets. Our results show that the combination of QCNN and federated averaging algorithm can be practicable, as the federated averaging algorithm trains the quantum algorithm in fewer communication training rounds. Distributed training by repetitively averaging local quantum nodes on COVID-19 and MedNIST datasets has been performed, where the data is distributed in a non-independent and identical, and real-world manner. The federated QCNN takes fewer communication rounds for training on increasing the client fraction in Non-IID1 distribution. The number of communication rounds taken by FedQCNN and FedCQNN to achieve target set accuracy are more in real-world distribution as compared to the Non-IID2 partitioned MedNIST dataset.

Moreover, the quantum federated models with the averaging algorithm converge to a higher testing accuracy in less communication rounds as compared with federated SGD model. The findings demonstrated the resilience and consistency of the FedQCNN and FedCQNN models when confronted with inadequate and unevenly distributed training data. While distributed quantum learning, the communication cost is the main constraint and it can be reduced by using small batch sizes or by adding more local SGD updates. The federated quanvolutional algorithm can help distribute computational resources on NISQ devices and preserve privacy. It would be extremely beneficial to deploy distributed quantum algorithms for enabling

scalable and privacy-preserving intelligent applications. The federated hybrid models are expected to be effective and robust on complex and high-dimensional datasets. In the future, with the availability of appropriate quantum computing resources, the proposed QFL framework can be implemented on high-dimensional images. It is an interesting fact that, if there is no prior information about the distribution of datasets, then finding out the distribution settings on the client side is a complex problem.

Data availability statement

Datasets are publicly available and referenced in the manuscript. The results and findings code are available upon request. The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Amandeep Singh Bhatia ⬤ https://orcid.org/0000-0002-4891-6160
Sabre Kais ⬤ https://orcid.org/0000-0003-0574-5346
Muhammad Ashraful Alam ⬤ https://orcid.org/0000-0001-8775-6043

References

[1] Rieke N *et al* 2020 The future of digital health with federated learning *npj Digit. Med.* **3** 1–7
[2] McMahan B, Moore E, Ramage D, Hampson S and Arcas B 2017 Communication-efficient learning of deep networks from decentralized data *Artif. Intell. Stat.* **54** 1273–82
[3] Rahman M, Hossain M, Islam M, Alrajeh N and Muhammad G 2020 Secure and provenance enhanced internet of health things framework: a blockchain managed federated learning approach *IEEE Access.* **8** 205071–87
[4] Zheng Z, Zhou Y, Sun Y, Wang Z, Liu B and Li K 2022 Applications of federated learning in smart cities: recent advances, taxonomy and open challenges *Connect. Sci.* **34** 1–28
[5] Xu J, Glicksberg B, Su C, Walker P, Bian J and Wang F 2021 Federated learning for healthcare informatics *J. Healthcare Inform. Res.* **5** 1–19
[6] Al Aziz M, Anjum M, Mohammed N and Jiang X 2022 Generalized genomic data sharing for differentially private federated learning *J. Biomed. Inform.* **132** 104113
[7] Tran N, Bao W, Zomaya A, Nguyen M and Hong C 2019 Federated learning over wireless networks: optimization model design and analysis *IEEE INFOCOM 2019-IEEE Conf. on Computer Communications* pp 1387–95
[8] Sprague M, Jalalirad A, Scavuzzo M, Capota C, Neun M, Do L and Kopp M 2018 Asynchronous federated learning for geospatial applications *Joint European Conf. on Machine Learning And Knowledge Discovery in Databases* pp 21–28
[9] Li Q, Wen Z, Wu Z, Hu S, Wang N, Li Y, Liu X and He B 2021 A survey on federated learning systems: vision, hype and reality for data privacy and protection *IEEE Trans. Knowl. Data Eng.* **35** 3347–66
[10] Bharti K *et al* 2022 Noisy intermediate-scale quantum algorithms *Rev. Mod. Phys.* **94** 015004
[11] Huang H, Broughton M, Mohseni M, Babbush R, Boixo S, Neven H and McClean J 2021 Power of data in quantum machine learning *Nat. Commun.* **12** 1–9
[12] Ciliberto C, Herbster M, Ialongo A, Pontil M, Rocchetto A, Severini S and Wossnig L 2018 Quantum machine learning: a classical perspective *Proc. R. Soc. A* **474** 20170551
[13] Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N and Lloyd S 2017 Quantum machine learning *Nature* **549** 195–202
[14] Bhatia A, Saggi M, Kumar A and Jain S 2019 Matrix product state-based quantum classifier *Neural Comput.* **31** 1499–517
[15] Sen P, Bhatia A, Bhangu K and Elbeltagi A 2022 Variational quantum classifiers through the lens of the Hessian *PLoS One* **17** e0262346
[16] Schuld M, Sinayskiy I and Petruccione F 2014 The quest for a quantum neural network *Quantum Inf. Process.* **13** 2567–86
[17] Abbas A, Sutter D, Zoufal C, Lucchi A, Figalli A and Woerner S 2021 The power of quantum neural networks *Nat. Comput. Sci.* **1** 403–9
[18] Xia R and Kais S 2018 Quantum machine learning for electronic structure calculations *Nat. Commun.* **9** 1–6
[19] Sajjan M, Li J, Selvarajan R, Sureshbabu S, Kale S, Gupta R, Singh V and Kais S 2022 Quantum machine learning for chemistry and physics *Chem. Soc. Rev.* **51** 6475–573
[20] Bhatia A, Saggi M and Kais S 2023 Quantum machine learning predicting ADME-Tox properties in drug discovery *J. Chem. Inf. Model.* (https://doi.org/10.1021/acs.jcim.3c01079)
[21] Schuld M and Killoran N 2019 Quantum machine learning in feature Hilbert spaces *Phys. Rev. Lett.* **122** 040504
[22] Kjaergaard M, Schwartz M, Braumüller J, Krantz P, Wang J, Gustavsson S and Oliver W 2020 Superconducting qubits: current state of play *Annu. Rev. Condens. Matter Phys.* **11** 369–95
[23] Henderson M, Shakya S, Pradhan S and Cook T 2020 Quanvolutional neural networks: powering image recognition with quantum circuits *Quantum Mach. Intell.* **2** 1–9
[24] Mattern D, Martyniuk D, Willems H, Bergmann F and Paschke A 2021 Variational quanvolutional neural networks with enhanced image encoding (arXiv:2106.07327)
[25] Yang X, He X, Zhao J, Zhang Y, Zhang S and Xie P 2020 COVID-CT-dataset: a CT scan dataset about COVID-19 (arXiv:2003.13865)
[26] Medical MNIST dataset 2022 (available at: www.kaggle.com/datasets/andrewmvd/medical-mnist)
[27] Chen S and Yoo S 2021 Federated quantum machine learning *Entropy* **23** 460
[28] Li W, Lu S and Deng D 2021 Quantum federated learning through blind quantum computing *Sci. China Phys. Mech. Astron.* **64** 1–8
[29] Xia Q and Li Q 2021 Quantumfed: a federated learning framework for collaborative quantum training *2021 IEEE Global Communications Conf. (GLOBECOM)* pp 1–6

[30] Yamany W, Moustafa N and Turnbull B 2021 OQFL: an optimized quantum-based federated learning framework for defending against adversarial attacks in intelligent transportation systems *IEEE Trans. Intell. Transp. Syst.* **24** 893–903

[31] Chehimi M and Saad W 2022 Quantum federated learning with quantum data *ICASSP 2022–2022 IEEE Int. Conf. on Acoustics, Speech And Signal Processing (ICASSP)* pp 8617–21

[32] Huang R, Tan X and Xu Q 2022 Quantum federated learning with decentralized data *IEEE J. Sel. Top. Quantum Electron.* **28** 1–10

[33] Zhang Y, Zhang C, Zhang C, Fan L, Zeng B and Yang Q 2022 Federated learning with quantum secure aggregation (arXiv:2207.07444)

[34] Kais S, Bhatia A and Alam M 2023 Quantum federated learning in healthcare: the shift from development to deployment and from models to data (https://doi.org/10.21203/rs.3.rs-2723753/v1)

[35] Bhatia A, Kais S and Alam M 2023 Handling privacy-sensitive clinical data with federated quantum machine learning *Bull. Am. Phys. Soc.*

[36] Broughton M, Verdon G, McCourt T, Martinez A, Yoo J, Isakov S, Massey P, Halavati R, Niu M and Zlokapa A 2020 and others tensorflow quantum: a software framework for quantum machine learning (arXiv:2003.02989)

[37] Cirq Developers 2022 Cirq (available at: https://doi.org/10.5281/zenodo.5182845)