Leveraging Quantum Contrastive Learning and Federated Averaging for Robust Out-of-Distribution Generalization

Asitha Kottahachchi Kankanamge Don[1], **and Ibrahim Khalil**[1]
1. STEM College, RMIT University, Melbourne, Australia
Correspondence: asitha.kottahachchi.kankanamge.don@student.rmit.edu.au

Introduction

Quantum Contrastive Learning (QCL) integrated with Vertical Federated Learning (VFL) offers a transformative solution to challenges in data scarcity, adversarial robustness, and scalability. The framework combines quantum-enhanced representations and supervised contrastive loss to ensure secure, scalable, and privacy-preserving distributed training. Applications span privacy-sensitive domains such as healthcare, finance, and sensor networks, where robust collaborative learning is essential without compromising data privacy.

Problem Definition

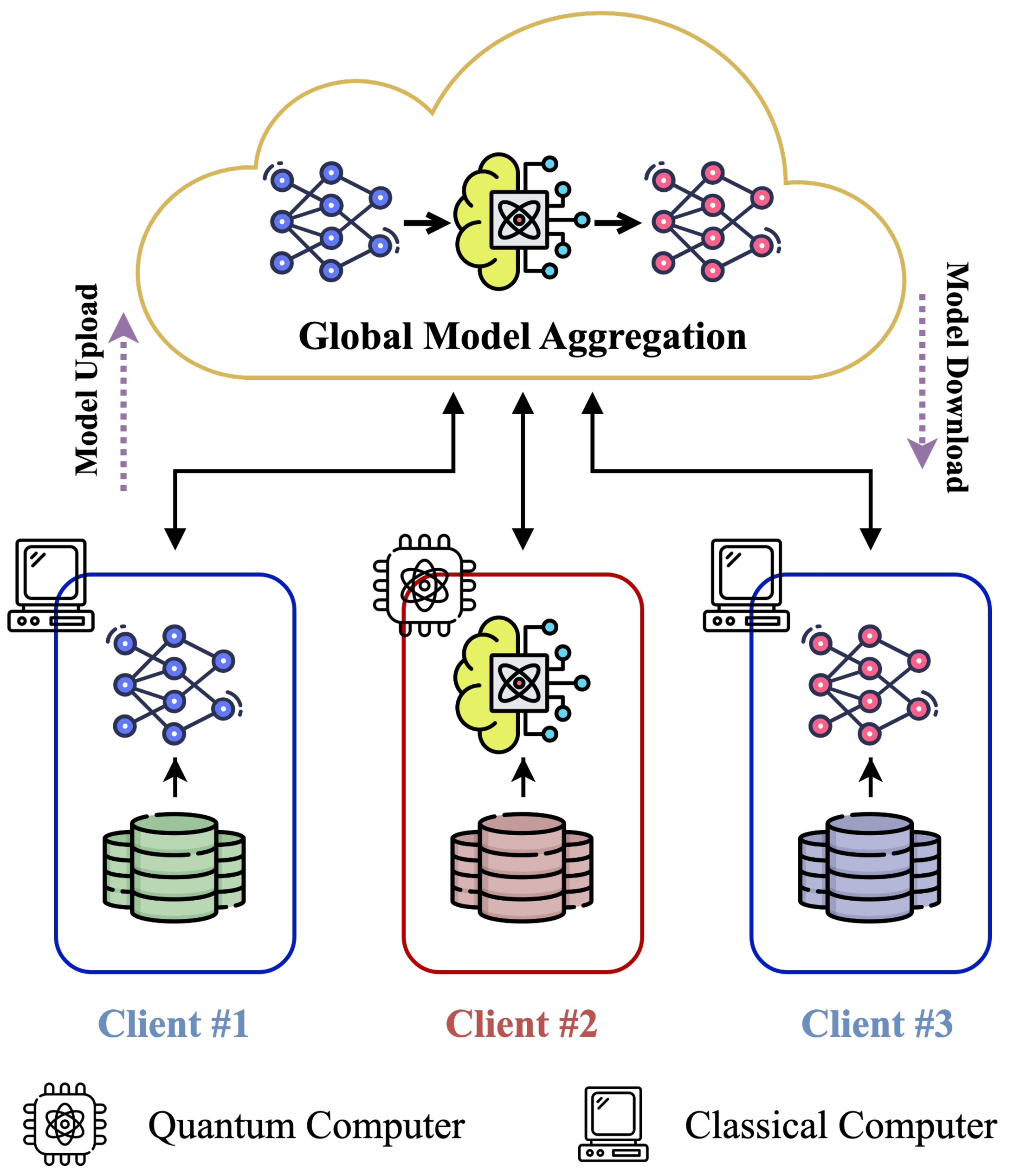

Figure 1: Federated Learning with Hybrid Clients: Quantum (Client #2) and classical clients (Client #1, Client #3) collaborate in global model aggregation and face challenges.

Traditional machine learning frameworks face significant obstacles in distributed and privacy-sensitive environments. Data scarcity limits model generalization in critical domains like medical imaging. Classical federated learning models are vulnerable to attacks, including membership inference, gradient leakage, and model inversion, exposing sensitive data. Moreover, these approaches often struggle to handle heterogeneous data distributions effectively. This work addresses these issues by integrating QCL and VFL into a secure, efficient, and scalable hybrid quantum-classical framework.

Framework Overview

The framework incorporates a Quantum Data Preprocessor to enhance feature diversity, followed by a Quantum Encoder that compresses input data into meaningful embeddings using parameterized quantum circuits. A Quantum Projection Head optimizes these embeddings for contrastive learning, ensuring robust feature separability. Enhanced Federated Averaging aggregates model updates across distributed nodes, leveraging quantum-aware weighting to prioritize nodes with richer feature representations.

Mathematical Framework

This research provides a theoretical analysis of the Quantum Contrastive Learning (QCL) framework with Vertical Federated Learning (VFL), deriving convergence guarantees under standard assumptions in quantum and federated learning. Empirical validation is planned to assess practical performance. The QCL framework's core comprises three components: Quantum Image Processor (Data Augmentation Module), Quantum Autoencoder, and Quantum Classifier, which together generate secure and robust embeddings for federated learning.

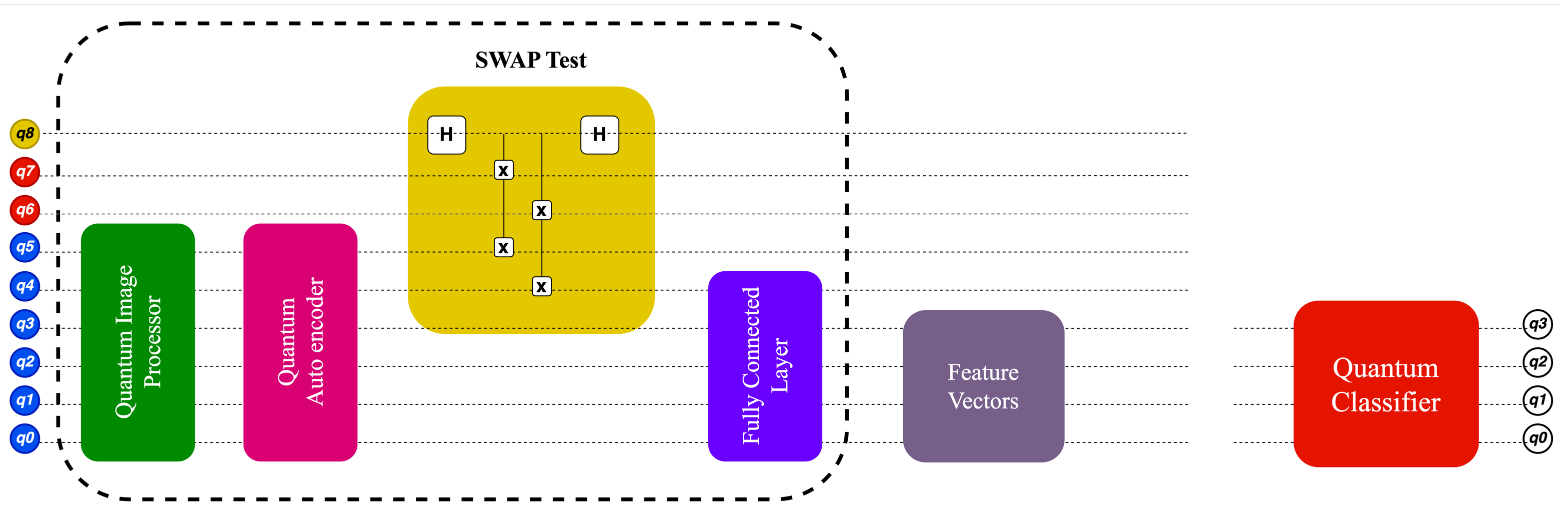

Figure 2: Quantum Contrastive Learning Framework: The framework integrates a quantum encoder, projection head, and a quantum classifier.

Quantum Image Processor. The Quantum Image Processor, also referred to as the Quantum Data Augmentation Module, applies unitary transformations to increase the diversity of input data. Given a classical image $X \in \mathbb{R}^{d\times d}$, it is transformed into a quantum state:

$$\psi_X = U(X)0, \tag{1}$$

where $U(X)$ is a parameterized quantum circuit that maps the classical image to a quantum state. This transformation ensures higher variability in the input space, which is essential for contrastive learning. The augmented quantum states are then passed to the Quantum Autoencoder.

Quantum Autoencoder and Projection Head. The Quantum Autoencoder compresses the high-dimensional quantum state into a lower-dimensional latent space. For a given quantum state ψ_X, the encoder performs the transformation:

$$\phi_X = V_{\text{enc}}\psi_X, \tag{2}$$

where V_{enc} is a parameterized quantum circuit for encoding. The goal is to learn a latent representation that minimizes the supervised contrastive loss, defined as:

$$\mathcal{L}_{\text{contrastive}} = -\frac{1}{|P(i)|}\sum_{i\in I}\sum_{p\in P(i)} \log \frac{\exp(h(X_i)\cdot h(X_p)/\tau)}{\sum_{a\in A(i)}\exp(h(X_i)\cdot h(X_a)/\tau)}, \tag{3}$$

where $h(X)$ represents the embeddings produced by the projection head, τ is the temperature parameter, $P(i)$ denotes positive pairs, and $A(i)$ represents all non-anchor pairs. The projection head ensures that embeddings of similar data points are aligned while those of different data points are well-separated.
The Quantum Autoencoder and Projection Head are trained jointly to optimize $\mathcal{L}_{\text{contrastive}}$, enabling the extraction of meaningful and robust quantum feature embeddings.

Quantum Classifier. After the Quantum Autoencoder and Projection Head are trained, the encoder network V_{enc} is frozen and reused to produce embeddings for the Quantum Classifier. The classifier operates on the latent states ϕ_X and maps them to classical feature vectors $h(X)$ through measurements:

$$h(X) = M\phi_X, \tag{4}$$

where M is the measurement operator. The Quantum Classifier is trained separately using a cross-entropy loss:

$$\mathcal{L}_{\text{classification}} = -\frac{1}{N}\sum_{i=1}^{N} y_i \log(\hat{y}_i), \tag{5}$$

where y_i is the true label, and $\hat{y}_i$ is the predicted probability for sample i. This two-stage training process ensures that the Quantum Classifier leverages the robust embeddings learned during contrastive training.

Global Framework Integration. The outputs of the Quantum Classifier are integrated into the VFL framework through the enhanced Federated Averaging algorithm. The global loss combines task-specific and regularization objectives:

$$\mathcal{L}(\Theta) = \frac{1}{N}\sum_{i=1}^{N} f(\Theta; X_i, y_i) + \lambda\sum_{k=1}^{K}\gamma(\Theta_k), \tag{6}$$

where $f(\Theta; X_i, y_i)$ is the classification loss, and $\gamma(\Theta_k)$ regularizes model parameters. Gradient updates are aggregated dynamically to avoid biases between quantum and classical clients. The aggregation weight w_k for client k is defined as:

$$w_k = \alpha\cdot\frac{\|H_k\|}{\sum_{j\in\text{quantum}}\|H_j\|} + (1-\alpha)\cdot\frac{\Delta\mathcal{L}_k}{\sum_{j=1}^{K}\Delta\mathcal{L}_j}, \tag{7}$$

where: - $\|H_k\|$: Norm of the quantum embeddings for quantum clients. - $\Delta\mathcal{L}_k$: Improvement in the global loss contributed by client k, for both quantum and classical clients. - α: A tunable parameter balancing contributions from embedding strength and loss improvement.
The dynamic adjustment ensures fairness by weighting nodes based on their contribution to global performance and representation richness. This approach minimizes biases and balances the influence of quantum and classical clients.

Convergence Analysis

The convergence of the Quantum Contrastive Learning (QCL) framework integrated with Vertical Federated Learning (VFL) is established under standard assumptions of Lipschitz continuity and bounded gradient variance.
Let $\mathcal{L}(\Theta)$ be the global loss function, assumed to be L-smooth:

$$\|\nabla\mathcal{L}(\Theta_1) - \nabla\mathcal{L}(\Theta_2)\| \le L\|\Theta_1 - \Theta_2\|. \tag{1}$$

The variance of local gradients across nodes is bounded:

$$\mathbb{E}\|\nabla\mathcal{L}_k(\Theta) - \nabla\mathcal{L}(\Theta)\|^2 \le \sigma^2. \tag{2}$$

The global update rule, governed by Federated Averaging, ensures:

$$\Theta_{t+1} = \Theta_t - \eta\sum_{k=1}^{K} w_k\nabla\mathcal{L}_k(\Theta_k), \tag{3}$$

where:

$$w_k = \frac{\|H_k\|}{\sum_{j=1}^{K}\|H_j\|}. \tag{4}$$

Under these assumptions, the framework achieves a convergence rate of:

$$\mathbb{E}[\mathcal{L}(\Theta_t)] - \mathcal{L}^* \le \mathcal{O}\left(\frac{1}{t}\right), \tag{5}$$

where $\mathcal{L}^*$ is the global minimum.
The quantum encoder reduces parameter dimensionality by compressing input states ψ_X into latent states ϕ_X, minimizing gradient variance and improving stability in heterogeneous environments.
Dynamic weighting ensures nodes with richer quantum embeddings contribute more to global updates, enhancing fairness and robustness in federated learning.

Summary

The proposed hybrid quantum-classical framework addresses critical challenges in distributed training by leveraging quantum-enhanced representations and robust privacy-preserving mechanisms. Its scalability and theoretical guarantees make it an ideal solution for privacy-sensitive applications, including healthcare, finance, and sensor networks.

Acknowledgments

This work is supported by the Australian Research Council Discovery Project (DP210102761).
We would like to extend our acknowledgment to Robert Shen from RACE (RMIT AWS Cloud Supercomputing Hub) and Jeff Paine from Amazon Web Services (AWS) for their invaluable provision of computing resources.