Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

Performance Analysis and Noise Impact of a Novel Quantum KNN Algorithm for Machine Learning

ASIF AKHTAB RONGGON[1], MD. SAIFUR RAHMAN[2]

[1]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology (e-mail: asifaftab172@gmail.com)

[2]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology (e-mail: saifur@eee.buet.ac.bd)

Corresponding author: Asif Akhtab Ronggon (e-mail: asifaftab172@gmail.com).

ABSTRACT This paper introduces a novel quantum K-nearest neighbors (QKNN) algorithm, which offers improved performance over the classical KNN technique by incorporating quantum computing techniques to enhance the classification accuracy, scalability, and robustness. The algorithm optimizes quantum data encoding by making use of Hadamard and rotation gates and enables efficient embedding of classical data into quantum states. Here, feature extraction is strengthened by entangling gates such as IsingXY and CNOT, which offer complex feature interactions and improved classification ability. To calculate the similarity measures between the various quantum states, a novel quantum distance metric is introduced, which is specifically designed based on the results of the swap test. This offers superior accuracy and computational efficiency compared to the traditional Euclidean distance metrics. We evaluated the proposed QKNN algorithm on three benchmark datasets, namely Wisconsin Breast Cancer, Iris, and Bank Note Authentication, and observed its superior performance relative to classical KNN (CKNN), Quantum Neural Network (QNN), and recent QKNN studies. The proposed QKNN algorithm is found to achieve prediction accuracies of 98.25%, 100%, and 99.27%, respectively, for the three datasets, while the QNN shows prediction accuracies of only 97.17%, 83.33%, and 86.18%, respectively. Moreover, the quantum noise challenges have been addressed by integrating a repetition code-based error mitigation strategy, which ensures stability of the algorithm and makes it resilient to noisy quantum environments. The results indicate that the proposed QKNN algorithm is scalable, more efficient, and robust. Thus, it has great potential for applications in complex and high-dimensional datasets, such as medical diagnostics, financial fraud detection, and security, where traditional methods often fail.

INDEX TERMS Quantum machine learning, quantum K-nearest neighbors, quantum distance computation, quantum feature mapping, quantum noise and error correction, classical KNN, hybrid classical-quantum systems and quantum neural network.

I. INTRODUCTION

The rapid advancements in quantum computing (QC) have spurred significant interest in developing algorithms that possess the potential to outperform their classical counterparts across various domains, including machine learning (ML). Quantum Machine Learning (QML) is a promising field of research that integrates quantum computational principles with traditional ML models to improve computational efficiency and scalability. Considerable attention has been directed toward quantum implementations of well-established classical algorithms, including the quantum K-nearest neighbor (QKNN) algorithm. They have attracted significant attention due to their rapid processing of high-dimensional data and greater computational efficiency than their classical counterparts [1].

On the other hand, the classical K-nearest neighbor (CKNN) is a widely used nonparametric ML approach that involves data mining and pattern recognition. Although it is a simple and efficient algorithm, it encounters significant computational hurdles, particularly for large datasets. This approach is computationally expensive for real-time applications because it relies on brute-force computation to find the closest neighbors [2]. Therefore, it became necessary to find other ways to accelerate the classification process, as the performance of the CKNN decreases significantly as the dataset size increases. One possible solution to these challenges is

QC, which can achieve exponential speedups compared to its classical counterpart. According to Wiebe [3], this is the cause for the introduction of a quantum variant known as the QKNN algorithm as a substitute. To enhance the effectiveness of searching neighbors and computing the distance of target data point from the neighbor data point, the QKNN algorithm utilizes quantum features such as superposition, entanglement, and quantum parallelism. In particular, quantum algorithms may offer more accurate similarity measurements compared to classical algorithms, and quantum systems can store large amounts of data [4]. For solving high-dimensional classification problems, the QKNN approach is a highly desirable option, as it uses quantum circuits to achieve exponential speedup compared to the traditional CKNN approach in specific applications [5].

Despite its theoretical advantages, the present status of quantum technology continues to impede the practical implementation of the QKNN approach. The full realization of QML algorithms is not yet fully realizable because of decoherence, noise, and the limited number of qubit counts of currently available quantum hardware. Furthermore, the development of an efficient quantum feature mapping technique poses a fundamental challenge. Therefore, research must be carried out to develop faster and more efficient techniques for the mapping process. Before practical implementation of the proposed QKNN algorithm, it is very important to evaluate its computational complexity, resilience to quantum noise, and classification accuracy in the presence of realistic quantum errors [6].

There are a few recent studies on the development of the QKNN approach that need special mention here. They are by Feng et al. [7], Maldonado et al. [8], Li et al. [9], and Bhaskaran et al. [10]. Feng et al. employed quantum amplitude encoding alongside a quantum polar distance metric, integrated with classical feature preprocessing, to enhance the performance of their QKNN model. Maldonado et al. propose a QKNN methodology using angle encoding through Quantum Random Access Memory (QRAM), combined with Grover's algorithm and the SWAP test for efficient quantum state comparison and neighbor identification. Li et al. present a QKNN methodology that simultaneously quantumizes both neighbor selection and K-value optimization using a novel quantum circuit based on quantum phase estimation, controlled rotation, and inverse phase estimation techniques. On the other hand, Bhaskaran et al. employ a QKNN methodology utilizing the Canberra distance metric, conducting a comparative analysis against CKNN across multiple datasets and K-values to evaluate classification performance under quantum enhancement.

Still, there are ample scopes to improve the accuracy of the QKNN algorithm. In this work, we propose a novel QKNN classifier that introduces several methodological advances not addressed in prior studies. The model employs Hadamard and R_z rotation gates for quantum state encoding, followed by a dedicated feature extraction layer utilizing IsingXY and CNOT gates to enhance entanglement and representation learning. To improve neighbor selection, we replace the conventional Grover search with a quantum SWAP test-based sorting mechanism, offering a more scalable and structured alternative. Furthermore, the framework integrates repetition code-based noise mitigation, a critical component lacking in existing QKNN architectures.

The research questions were carefully chosen to address key challenges that hinder the practical implementation of QKNN algorithms, including the effects of quantum noise, limited qubit resources, and hardware constraints. By employing performance analysis, noise resilience assessment, and hardware-aware design, this study aims to bridge the gap between the theoretical promise and practical implementation of the proposed QKNN on near-term quantum devices. Consequently, the objectives formulated here closely align with the research questions and identified gaps, ensuring a coherent and focused investigation. This alignment establishes a cohesive framework that guides the study from its motivation to its execution. The major contributions of the current research are as follows:

- Propose a novel QKNN algorithm that combines efficient quantum data encoding with entangling gates and a distance metric based on the swap test to improve classification accuracy and efficiency.
- Introduce a conceptual framework that uses quantum superposition and entanglement to obtain richer feature representation and robust similarity measurement, balancing expressivity, noise resilience, and hardware efficiency.
- Develop an integrated noise model with Pauli error channels and suggest a repetition code–based error mitigation to improve the robustness of QKNN on noisy quantum hardware.
- Provide comprehensive benchmarking against CKNN, QNN, and recent QKNN studies across standard datasets, demonstrating superior performance, scalability, and noise tolerance.

Here is how the rest of the paper is organized. In Section II, a comprehensive review of the literature on QML, QKNN, and QNN is included. The description of the dataset, algorithmic specifics, and experimental setup is presented in Section III. Classification accuracy, computational efficiency, and noise resistance were compared between traditional CKNN and QKNN in Section IV, which also covers the results of the performance evaluation. Section V includes the conclusions and findings of this research and suggests further research directions in QML. For clarity, a summary of the complex symbols and notation used throughout this document is provided in Table 1.

II. RELATED WORK

QML is a novel concept that optimizes the traditional ML algorithms using quantum physics. New QC capabilities have shown that quantum algorithms can uniquely solve conventional computational problems, including data classification, feature selection, and model optimization. The versatility and

TABLE 1: Nomenclature

Symbol	Description	Unit / Note
$D(\psi_{\text{train},i}, \psi_{\text{test}})$	Quantum distance metric based on swap test between training and test states	Dimensionless
$\mathcal{E}(\rho)$	Quantum noise channel acting on density matrix ρ	—
H	Hadamard gate operator	Unitary operator
$R_Z(\theta)$	Rotation gate about Pauli-Z axis parameterized by θ	Unitary operator
$R_Y(\theta)$	Rotation gate about Pauli-Y axis parameterized by θ	Unitary operator
S_i	Stabilizer operator for error detection in repetition code	Hermitian operator
U	General unitary operator for combined quantum transformations	Unitary operator
$U_{\text{embed}}(x)$	Quantum embedding operator encoding classical data x	Unitary operator
$U_{\text{entangle}}(\theta)$	Unitary operator implementing parameterized entangling gates	Unitary operator
$U_{\text{IsingXY}}(\theta)$	Two-qubit IsingXY interaction gate parameterized by θ	Unitary operator
X, Y, Z	Pauli operators for bit flip, bit-phase-flip, and phase flip errors	Hermitian operators
ρ	Density matrix representing the quantum state	Operator
$P_{\text{logical error rate}}$	Logical error probability after error correction	Probability
p	Physical qubit error probability	Probability
$\mathbb{I}(\cdot)$	Indicator function counting class occurrences (1 if true, else 0)	—

effectiveness with which the QKNN and QNNs categorize large datasets have brought them to the forefront among existing QML algorithms.

In recent years, an upsurge of interest in the integration of QC principles with ML algorithms, especially the QKNN and QNN classifiers, has been noticed. For large, high-dimensional datasets and multiclassification jobs, the objective of such algorithms is to improve both the computational efficiency and classification accuracy simultaneously. Despite the encouraging findings demonstrated by quantum techniques, there are still several scaling problems and realistic application challenges that persist, particularly in relation to quantum noise and decoherence [11].

An essential part of QML is quantum feature encoding, which maps classical input to their equivalent quantum states. Several encoding methods, including amplitude encoding, angle encoding, and quantum random access coding (QRAC), have been examined for the purpose of effectively representing the classical information into the quantum systems [12] [13]. Nguyen et al. [14] conducted a review of quantum visual encoding strategies and identified the Quantum Information Gap (QIG), which results in information loss between the classical and quantum features. The Quantum Information Preserving (QIP) loss function was proposed to reduce this gap, thus improving the effectiveness of QML algorithms. Sakka et al. [15] presented an agentic framework capable of autonomously generating, evaluating, and refining quantum feature maps through the use of large language models. This method revealed that the algorithm was worthy of identifying feature maps that surpassed current quantum baselines and achieved a comparable accuracy relative to classical kernels in datasets such as MNIST.

Numerous studies have investigated the QKNN algorithms, demonstrating potential improvements over the CKNN methods. Jarir and Quafafou [16] proposed a hybrid quantum-classical KNN approach for text classification, which utilizes quantum algorithms for feature extraction. However, their study did not provide direct accuracy comparisons with classical methods. Zardini et al. [17] introduced a QKNN algorithm based on Euclidean distance estimation, which improves computational efficiency in high-dimensional spaces. However, the study lacked explicit accuracy benchmarks. Li et al. [9] introduce a quantum K-nearest neighbors (QKNN) classifier that simultaneously quantumizes both neighbor selection and K value determination by formulating an objective function combining the least squares loss and sparse regularization using the HHL algorithm. The algorithm employs a quantum circuit using phase estimation, controlled rotation, and inverse phase estimation.

In addition to KNN, QNN has become a viable approach for enhancing the tasks of classification. Quantum-enhanced neural networks take advantage of QC to offer significant improvements in both speed and accuracy. Berti et al. [18] explore the transformation of classical K-nearest neighbor algorithms into their quantum equivalents, focusing on two distinct designs, namely, amplitude encoding and basis encoding. Highlights the impact of these encoding methods on algorithm structure, distance metrics, and performance while addressing challenges in data preparation and the theoretical advantages of quantum algorithms over classical ones. In their study, Wang et al. [19] applied quantum KNN to handwritten digit recognition, achieving 98% precision in the MNIST dataset, which outperformed the CKNN algorithms achieving an accuracy of around 94%.

Further developments in QML have seen significant progress in QNN. Bhaskaran and Prasanna [20] performed an accuracy analysis comparing the classical and quantum-enhanced KNN algorithms, reporting accuracy improvements of up to 18% in favor of quantum-enhanced methods. Feng et al. [7] proposed an enhanced quantum KNN classification algorithm using the quantum polar distance, which demonstrated significant improvements in classification accuracy, especially for high-dimensional data. Xiang et al. [21] explored hybrid quantum-classical convolutional neural networks (QCCNNs) for breast cancer diagnosis, demonstrating improvements in both accuracy and computational efficiency over classical approaches. Zhul et al. [22] propose a fast clustering approach that integrates graph-regularized non-negative matrix factorization with quantum clustering to harness the strengths of both, enhancing speed and accuracy through a quantum-inspired clustering algorithm that combines matrix factorization and quantum computing principles.

In summary, a good number of studies have focused on improving the computational efficiency and scalability of quantum KNN and QNN algorithms, some of which have also demonstrated notable improvements in classification accuracy. However, direct and consistent accuracy comparisons between quantum-enhanced and classical methods across various datasets remain an area for further research.

III. METHODOLOGY

This section outlines the methodology used to evaluate the performance of the proposed QKNN algorithm. We focus on data preprocessing, implementation of the QKNN algorithm, and comparison with CKNN and QNN. In addition, we discuss the impact of quantum noise on the classification accuracy and present strategies for error mitigation.

A. DATASET DESCRIPTION

For evaluating the proposed QKNN algorithm, three benchmark datasets are utilized: Wisconsin Breast Cancer, Iris, and Bank Note Authentication.

Wisconsin Breast Cancer Dataset: This dataset contains 569 instances, each described by 30 features derived from digitized images of fine needle aspirates (FNA) of breast masses. The task involves binary classification to differentiate between benign and malignant tumors [23].

Iris Dataset: Introduced by Fisher [24], the Iris dataset consists of 150 instances characterized by four features (sepal length, petal length, sepal width and petal width). The classification task distinguishes between three species of iris flowers: versicolor, setosa, and virginica. Its ease of use and neatly divided classes make it a standard benchmark for classification tasks.

Bank Note Authentication Dataset: This dataset comprises 1,372 instances, each with 4 features extracted from images of genuine and forged banknotes. The goal is a binary classification task to distinguish between authentic and counterfeit notes [25].

The basic attributes of the three datasets are shown in Table 2.

TABLE 2: The basic attributes of the three datasets.

Dataset	Instances	Features	Classes	Reference
Wisconsin Breast Cancer	569	30	2	[23]
Iris	150	4	3	[24]
Bank Note Authentication	1372	4	2	[25]

B. FEATURE SELECTION USING THE CHI-SQUARE TEST

Feature selection is a crucial step in reducing dimensionality, improving model efficiency, and preventing overfitting. In this study, we apply the chi-square (χ^2) test to assess the dependency between each feature and the target class.

The chi-square statistic for each feature F_i is calculated as:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

where O_i is the observed frequency, E_i is the expected frequency under the null hypothesis of independence, and k is the number of bins or categories.

For continuous features, discretization is applied to convert them into categorical values, with expected frequencies calculated relying on the notion that there is statistical independence between the feature and the target class.

A significant result (typically with a p-value less than 0.05) specifies that the feature is relevant to the target class and should be retained. Features with high p values, indicating weak association, are discarded to reduce model complexity and improve generalization [26]. This method has been successfully used in numerous ML applications to select relevant features and enhance model performance [27].

In this study, the UCI breast cancer dataset, originally comprising multiple features, is reduced to the four most statistically significant using the chi-square test for quantum encoding. The remaining two datasets naturally contain four features each. Standardizing the count of features across datasets ensures uniform dimensions of the quantum circuit, facilitating an equitable comparison of experimental results. Employing chi-square feature selection is critical in QML, where qubit availability is a fundamental limitation. Since each feature maps to one or more qubits, a higher dimensionality directly increases quantum resource demands and circuit complexity. By rigorously selecting only the most relevant features, the chi-square test minimizes qubit usage without compromising predictive accuracy. This dimensionality reduction not only enhances computational tractability on near-term quantum devices but also mitigates the noise and decoherence effects inherent to deeper circuits. Thus, chi-square-based feature selection provides a systematic and practical strategy to reconcile the informativeness of the dataset with the constraints of current quantum hardware.

C. QUANTUM MACHINE LEARNING

QML is an emerging paradigm that integrates the principles of QC with ML algorithms. The fundamental advantage of QML arises from quantum phenomena such as superposition, entanglement, and quantum parallelism, which enable certain computations to be performed more efficiently than their classical counterparts [6]. These quantum properties provide the potential for QML to outperform classical models, particularly in tasks that involve high-dimensional space, optimization, and large-scale data processing.

One of the key elements of QML is the encoding of quantum data, where the classical data is mapped to quantum states. Unlike classical representations, quantum encoding allows data to exist in a superposition of multiple states simultaneously, enabling parallel computation. This property is further enhanced by quantum feature maps, which transform classical inputs into quantum Hilbert spaces, allowing for more expressive data processing compared to classical kernels [4].

A popular approach in QML involves Variational Quantum Circuits (VQCs), which use quantum gates parameterized by tunable variables. These gates are optimized to learn patterns from the data. Hybrid quantum-classical models combine QC with classical optimization techniques, showing promising results in tasks such as classification, clustering, and generative modeling [28]. These hybrid models are particularly useful given the current limitations of noisy intermediate-scale quantum (NISQ) devices, which suffer from limited qubit coherence and gate fidelity.

QML is also advancing in quantum kernel methods and quantum support vector machines (QSVMs). Quantum-enhanced kernels facilitate higher-dimensional feature mappings that are often intractable classically, providing the ability to handle more complex data structures [4]. Furthermore, quantum Boltzmann machines and quantum generative adversarial networks (QGANs) have been explored for generative learning tasks such as image synthesis and anomaly detection [29].

Despite its promise, QML faces several challenges. For example, current quantum hardware is limited by factors such as qubit count, error rates, and coherence times, which restrict the practical implementation of large-scale quantum models. Furthermore, efficient quantum-specific optimization techniques and quantum circuit architectures remain active areas of research. One significant bottleneck is quantum data loading, where encoding large classical datasets into quantum states is a non-trivial task [30].

The integration of QC into ML has the potential for significant advances in both theoretical and practical applications. As quantum hardware improves, QML is expected to lead to breakthroughs in fields such as drug discovery, materials science, and financial modeling, where complex simulations require computational power beyond classical capabilities [1]. Ongoing research in quantum algorithms, noise mitigation, and quantum-classical hybrid frameworks will be crucial in unlocking the full potential of QML.

In conclusion, QML represents a transformative approach to ML with the potential to address computational bottlenecks that classical systems struggle with. As both experimental and theoretical developments progress, QML may become a powerful tool for solving problems that were previously considered intractable.

D. CLASSICAL VS. QUANTUM FEATURE SPACES

In classical ML, data is represented as feature vectors in an n-dimensional Euclidean space given as follows:

$$\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n \tag{2}$$

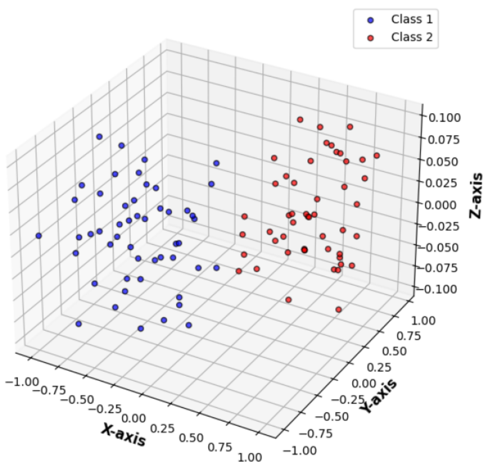
Algorithms, like SVMs, rely on kernel methods to map data into higher-dimensional spaces, allowing for separation of nonlinearly separable data. However, this mapping increases computational complexity [31].

In contrast, QC encodes data into quantum states via superposition and entanglement. A classical data point $\mathbf{x}$ is assigned to a quantum state represented as follows:

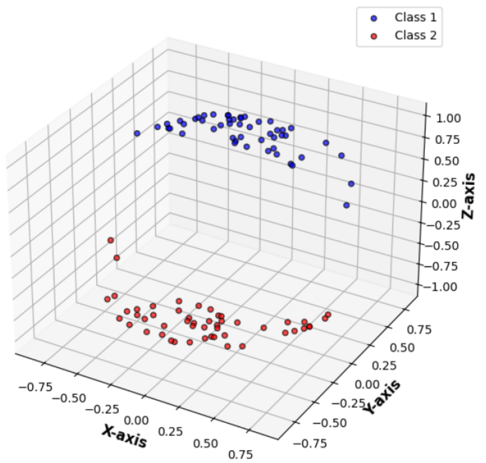
$$|\psi_x\rangle = U(x)\,|0\rangle \tag{3}$$

where $U(x)$ is a unitary operator and $|0\rangle$ represents the quantum state [4]. Quantum feature spaces inherently embed data into exponentially larger spaces, enhancing the ability of the model to capture complex patterns and relationships that are difficult to represent classically.

Quantum parallelism allows for the efficient exploration of high-dimensional spaces, which makes QML particularly suitable for large, complex datasets. Figure 1 shows classical and quantum feature spaces.



(a) Classical Feature Space



(b) Quantum Feature Space

FIGURE 1: Comparison between classical and quantum feature spaces.

From Figure 1(a), it is seen that the class 1 and class 2 data points are not separable, while Figure 1(b) shows that they are separable. In fact, the quantum encoding technique maps data into an exponentially large quantum Hilbert space. Thus, it offers a more powerful feature representation technique compared to the classical methods and leverages exponential dimensionality for better pattern separation.

E. QUANTUM GATES AND THEIR FUNCTIONS

Quantum gates are the fundamental building blocks of QC, analogous to classical logic gates, but they operate on quantum bits (qubits) rather than classical bits. Unlike classical gates, quantum gates are represented by unitary matrices, ensuring reversibility in quantum operations [32]. These gates manipulate quantum states by leveraging principles such as superposition and entanglement, enabling complex computations that classical computers cannot efficiently simulate.

Table 3 presents an overview of the commonly used quantum gates, their matrix representations, and a brief description of their functionalities.

In the construction of quantum algorithms such as Grover's algorithm for unstructured search and Shor's algorithm for integer factorization, quantum gates are fundamental [35] [36]. One common use of the Hadamard gate is the construction of superpositions which enable quantum parallelism. An essential part of quantum processing, the entanglement of qubits, requires the CNOT gate [33]. A key component of fault-tolerant QC design is the Toffoli gate, which is classically applicable to all reversible computations.

In quantum circuits, gates facilitate the precise manipulation of qubit states, thereby allowing for the execution of quantum algorithms that can exceed the performance of their classical counterparts. The flexibility of quantum computers depends on the effective deployment and enhancement of quantum gates as the hardware evolves.

F. THE PROPOSED QKNN ALGORITHM

The proposed QKNN algorithm introduces a novel approach to classification by integrating quantum computational principles into the conventional CKNN framework. This quantum-enhanced method capitalizes on quantum state encoding, feature extraction through quantum transformations, and quantum distance computations to improve classification accuracy and efficiency. By exploiting quantum-mechanical properties such as superposition and entanglement, QKNN has the potential to process complex datasets more effectively than the classical methods, particularly for high-dimensional and intricate feature spaces. Figure 2 represents the flow diagram of the proposed QKNN algorithm.

Quantum Data Encoding

The first step in the proposed QKNN algorithm involves encoding classical data points into quantum states, enabling the exploitation of quantum computation advantages. Given a training dataset $X_{\text{train}} = \{x_{\text{train}}^{(1)}, x_{\text{train}}^{(2)}, \ldots, x_{\text{train}}^{(N)}\}$ and a test dataset $X_{\text{test}} = \{x_{\text{test}}^{(1)}, x_{\text{test}}^{(2)}, \ldots\}$, each data point is assigned to a quantum state using Hadamard and rotation gates, ensuring efficient representation and accessibility within the quantum framework.

The Hadamard transformation initializes qubits into an equal superposition state, laying the foundation for parallel quantum computations. The states are given by as follows:

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \tag{4}$$

Applying Hadamard gates across d qubits produces a uniform superposition state that facilitates efficient quantum data processing. The states are represented by:

$$|\psi_{\text{init}}\rangle = \frac{1}{\sqrt{2^d}} \sum_{i=0}^{2^d-1} |i\rangle \tag{5}$$

where d is the dimension of the feature vector.

Rotation gates R_Z are applied to embed feature values into the quantum state. The rotation gate is specified as follows:

$$R_Z(\theta) = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix} \tag{6}$$

TABLE 3: Matrix Representation and Functions of the Common Quantum Gates.

Gate	Matrix Representation	Functionality Description
Hadamard (H)	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$	Creates equal superposition of $\vert 0\rangle$ and $\vert 1\rangle$ states [32].
Pauli-X (X)	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	Equivalent to the classical NOT gate; flips $\vert 0\rangle$ to $\vert 1\rangle$ and vice versa [32].
Pauli-Y (Y)	$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$	Rotates the qubit state around the Y-axis of the Bloch sphere by π radians [33].
Pauli-Z (Z)	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	Applies a phase shift of π to the $\vert 1\rangle$ state [34].
CNOT (CX)	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	A two-qubit gate that flips the target qubit if the control qubit is $\vert 1\rangle$ [32].
Toffoli (CCX)	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$	A controlled-controlled NOT gate used in reversible computing [35].
SWAP	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	Exchanges the quantum states of two qubits [32].
Phase (S)	$\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$	Applies a phase shift of $\pi/2$ to the $\vert 1\rangle$ state [34].
T-gate (T)	$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$	Introduces a phase shift of $\pi/4$, playing a key role in quantum universality [32].

where $\theta = \mathbf{2\pi x_i}$ and x_i is the normalized i-th feature value in the input vector. As a result, each classical data point is transformed into the quantum representation:

$$|x_{\text{train}}\rangle = \bigotimes_{i=1}^{d} R_Z(\theta_i) H |0\rangle \tag{7}$$

where $\bigotimes$ denotes the tensor product across qubits. This encoding capability enables the quantum system to process multiple data points, improving computational efficiency. Furthermore, the proposed quantum data encoding scheme serves as a fundamental conceptual advancement by embedding classical data into a high-dimensional quantum feature space that exploits quantum superposition and interference for enhanced expressivity.

Quantum Feature Extraction

Quantum transformations serve to enhance data separability by introducing higher-order correlations and structural modifications within the encoded quantum states. In our QKNN framework, the IsingXY gate is employed to establish entanglement between qubits, thereby enabling richer feature interactions. Formally, the IsingXY gate is defined as:

$$U_{\text{IsingXY}}(\theta) = \exp\left(-i\frac{\theta}{2}(\sigma_x^{(1)}\sigma_x^{(2)} + \sigma_y^{(1)}\sigma_y^{(2)})\right) \tag{8}$$

where $\sigma_{\mathbf{x}}^{(\mathbf{k})}$ and $\sigma_{\mathbf{y}}^{(\mathbf{k})}$ denote the Pauli X and Y operators acting on qubit k, respectively.

The use of this gate facilitates the generation of multipartite entanglement, which captures complex and nonlinear correlations among input features. This enhanced expressivity in the

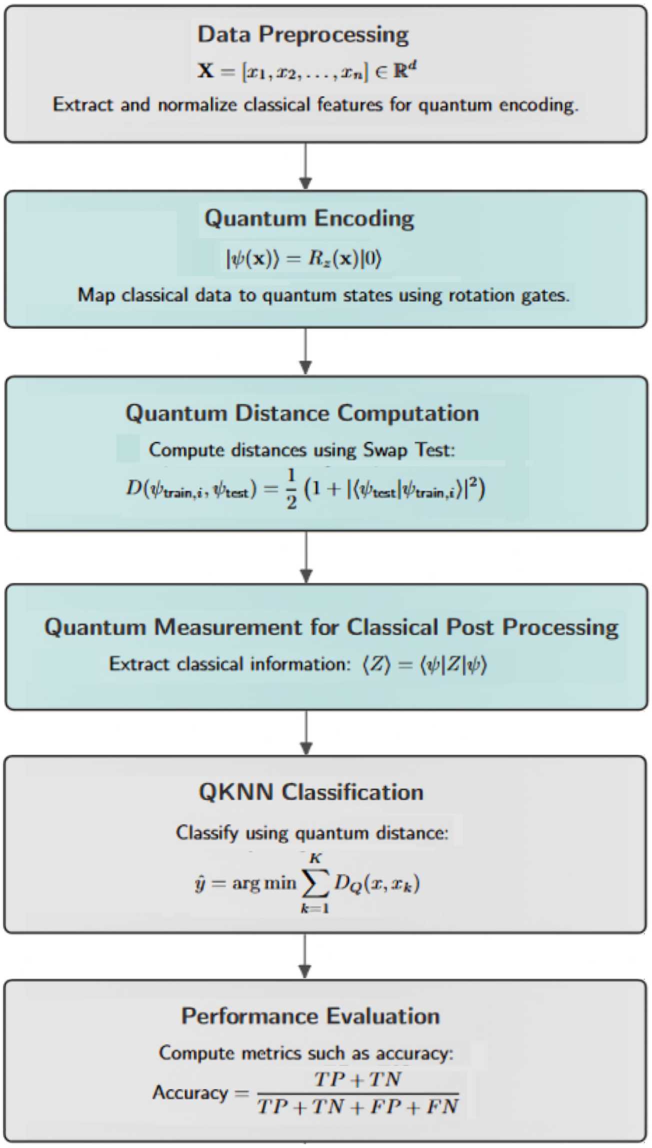


FIGURE 2: Flow diagram of the Proposed QKNN Algorithm

quantum feature space supports improved class separability that is difficult to achieve with classical transformations.

Additionally, Controlled-NOT (CNOT) gates are incorporated to further reduce class overlap by creating controlled entanglement patterns that promote feature space individuality among data points. The overall quantum state transformation can be expressed as:

$$|\psi_{\text{train}}\rangle \to U|\psi_{\text{train}}\rangle, \quad |\psi_{\text{test}}\rangle \to U|\psi_{\text{test}}\rangle \tag{9}$$

where U denotes the combined unitary operation of IsingXY and CNOT gates.

The deliberate selection of IsingXY and CNOT gates reflects a theoretical design choice aimed at balancing circuit expressivity and noise robustness. This strategic entanglement structure enables our QKNN algorithm to surpass CKNN by creating more discriminative representations of quantum features, thereby significantly enhancing classification performance.

Quantum Distance Calculation and Neighbor Selection

The similarity between a test quantum state $|\psi_{\text{test}}\rangle$ and each training quantum state $|\psi_{\text{train},i}\rangle$ is evaluated using a quantum distance metric based on the well-established swap test, defined as:

$$D(\psi_{\text{train},i}, \psi_{\text{test}}) = \frac{1}{2}\left(1 + |\langle\psi_{\text{test}}|\psi_{\text{train},i}\rangle|^2\right). \tag{10}$$

Although the swap test is a well-established quantum algorithmic technique for estimating the overlap between quantum states, our work introduces a novel integration of the swap test–based distance metric within a comprehensive and scalable QKNN framework.

This approach exploits the ability of the swap test to quantify subtle quantum features, including coherence and entanglement, within the exponentially large Hilbert space, thereby providing a similarity measure that transcends the limitations of classical distance metrics. Although the swap test itself is not new, our contribution lies in its rigorous theoretical justification and effective incorporation as a fundamental component of a noise-resilient quantum KNN classifier. This integration bridges the core principles of quantum computing with practical QML implementations optimized for near-term quantum hardware. The nearest neighbors for the test point are selected based on the smallest quantum distances, ensuring optimal classification.

Classification via Quantum Nearest Neighbors

After identifying the nearest neighbors, the classification is performed through a majority voting mechanism. The predicted class label is determined as follows:

$$C_{\text{pred}}(X_{\text{test}}) = \arg\max_{C_i} \sum_{i=1}^{K} \mathbb{I}(C_{\text{train},i} = C_i) \tag{11}$$

where $\mathbb{I}(\cdot)$ is an indicator function that counts occurrences of each class among the selected neighbors. The class with the highest frequency is assigned to the test instance, ensuring robust classification.

Figure 3 shows the proposed QKNN architecture. This architecture provides a novel approach to improve the efficiency of KNN classification by incorporating QC principles.

Rationale and Impact of Gate Selection

The selection of IsingXY, CNOT and SWAP gates in our QKNN algorithm reflects their complementary roles in enhancing algorithmic efficiency and hardware compatibility. IsingXY gates implement native two-qubit interactions from the XY-model Hamiltonian, common in superconducting qubits and trapped ions. Their use enables hardware-efficient ansatzes that reduce circuit depth and gate overhead compared to CNOT-only designs, improving expressivity and noise resilience.

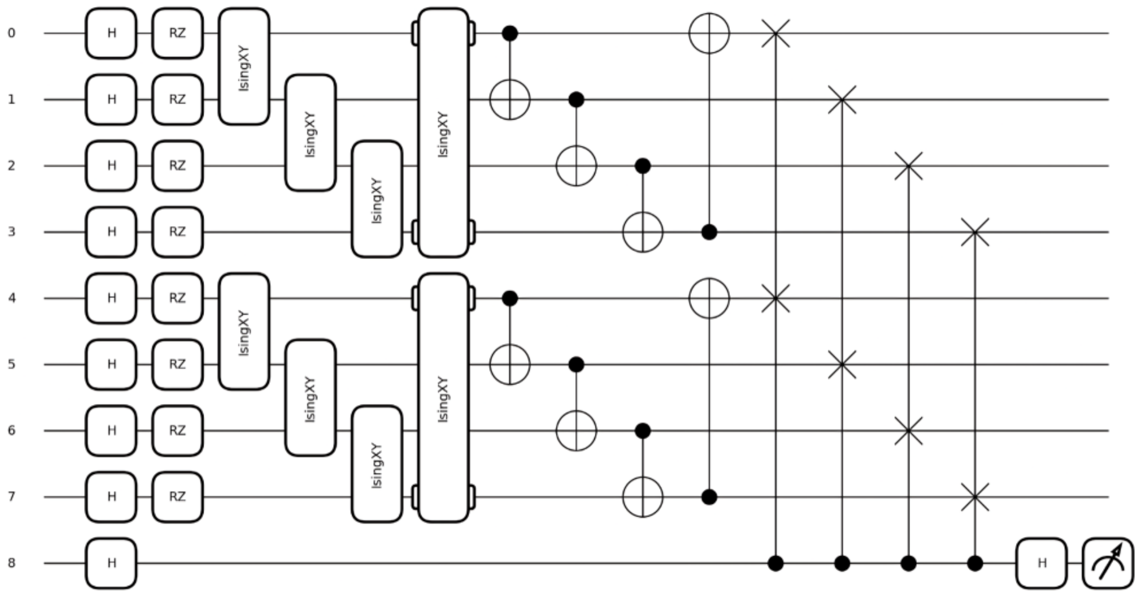


FIGURE 3: The proposed QKNN architecture.

Although both CNOT and SWAP gates are multiqubit operations that increase the depth and circuit complexity, they perform distinct and essential functions. CNOT gates enable feature extraction based on entanglement within the circuit, encoding complex data correlations critical to accurate classification. In contrast, SWAP gates primarily facilitate the swap test, a fundamental procedure for assessing quantum state similarity integral to distance estimation and classification in the KNN framework. Additionally, SWAP gates support qubit routing to address hardware connectivity limitations. Despite increasing circuit depth and compilation overhead, the complementary functions of the CNOT gates in feature processing and the SWAP gates in similarity evaluation justify their combined use.

Together, this set of gates strikes a balance between expressivity, noise robustness, and hardware adaptability, ensuring effective implementation of the QKNN algorithm in near-term quantum devices with diverse native gates and limited connectivity.

Description of the Proposed QKNN Algorithm

The steps performed in the proposed QKNN algorithm are included in Algorithm 1. To identify the nearest neighbors for classification, the proposed algorithm first takes quantum transformations as input for feature extraction and utilizes quantum distances determined from the swap test.

Table 4 presents a comparative analysis of various existing QKNN approaches with the proposed QKNN, focusing on key factors such as encoding methods, feature extraction techniques, distance metrics, neighbor selection, and noise mitigation strategies. Compared to existing methods, our proposed QKNN algorithm introduces significant performance improvements with a simplified quantum circuit and smaller circuit depth. Our method aims to optimize the performance and scalability of QKNN models, making them more feasible for real-world applications where computational resources and quantum hardware capabilities are constrained.

Algorithm 1 The Proposed Quantum K-Nearest Neighbors (QKNN) Algorithm

Require: Training dataset $X_{\text{train}}, Y_{\text{train}}$, test instance X_{test}.
Ensure: Predicted class label C_{pred}.
1: Encode classical data into quantum states using Hadamard and rotation gates.
2: Apply quantum transformations (IsingXY, CNOT) for feature extraction.
3: Compute quantum distances using the swap test.
4: Identify nearest neighbors based on minimum quantum distances.
5: Determine the class using majority voting among the selected neighbors.
6: **Return** C_{pred}.

By leveraging quantum-enhanced distance evaluation and feature transformations, the QKNN algorithm provides improved efficiency and performance in classification tasks. The integration of quantum mechanics in ML algorithms enables faster computations and enhanced pattern recognition, making QKNN a promising approach for ML applications in QC environments.

G. QUANTUM NEURAL NETWORK FOR CLASSIFICATION

Traditional QNNs combine the computational advantages of QC with the learning capabilities of classical neural networks. QNNs leverage quantum mechanical properties such as superposition, entanglement, and quantum interference to perform

TABLE 4: Comparison between Various Existing QKNN Approaches with the Proposed QKNN.

Study	Encoding Method	Feature Extraction	Distance Metric	Neighbor Selection	Noise Mitigation
Gao et al. [37]	Amplitude Encoding with Quantum Random Access Memory (QRAM)	Classical	Quantum Mahalanobis Distance	Grover search based on quantum distance	Not addressed
Zardini et al. [17]	Amplitude Encoding	Classical	Quantum Euclidean Distance	Quantum distance based sorting	Not addressed
Jing Li et al. [38]	Binary Encoding Method	Not addressed	Quantum Hamming Distance	Quantum distance based sorting	Not addressed
Feng et al. [7]	Amplitude Encoding	Classical	Quantum Polar Distance	Grover search based on quantum distance	Not addressed
Proposed Work	Hadamard + R_Z Rotation Gates	IsingXY + CNOT gates for entanglement	Swap test	Quantum swap test based sorting	Repetition Code

ML tasks efficiently [39] [40]. Unlike classical neural networks that rely on weighted connections between neurons, QNNs utilize parameterized quantum circuits (PQCs) to process and classify data [1] [41].

The fundamental building block of a QNN is a quantum circuit composed of unitary transformations. Given an input state $|\psi_{in}\rangle$, a QNN applies a series of quantum gates parameterized by θ:

$$|\psi_{out}\rangle = U(\theta)|\psi_{in}\rangle \tag{12}$$

where $U(\theta)$ is a parameterized unitary operator defined by the quantum circuit with trainable parameters θ [42]. The trainable parameters θ are optimized using classical gradient-based methods, often through a quantum-classical hybrid approach involving variational quantum eigensolver (VQE) or quantum natural gradient methods [43].

The capacity of QNNs to encode data into higher-dimensional Hilbert spaces is one of its most important characteristics; this allows for more detailed decision boundary representations than is possible with more conventional methods. The generalized achievement in ML tasks can be enhanced using this quantum enabled feature space [4].

QNNs use quantum circuits for data measurement, translation, and encoding, which is a new way to tackle classification challenges. This section outlines the theoretical foundations and methodology employed in the proposed quantum classification structure. The QNN architecture with parameterized quantum circuits is illustrated in Figure 4.

Quantum Encoding and Circuit Architecture

Given a dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$, where $x_i \in \mathbb{R}^d$ represents the feature vector and y_i denotes the corresponding label, the QNN begins with the preparation of the quantum state. Classical data is embedded into a quantum state via angle embedding as follows:

$$U_{embed}(x) = \prod_{i=1}^{d} R_Y(x_i)|0\rangle \tag{13}$$

where $R_Y(x_i) = e^{-ix_iY/2}$ denotes a rotation by angle x_i about the Pauli-Y axis acting on the i-th qubit, and $\mathbf{x} = [x_1, x_2, \ldots, x_d]^\top$ is the classical input feature vector [32]. Importantly, these local rotations act independently on each qubit and do not generate entanglement.

To introduce nonlinearity and entanglement, we employ PennyLane's strongly entangling layer, which consists of parameterized single-qubit R_Z rotations followed by CNOT gates arranged in a nearest-neighbor ring topology. The CNOT gates establish pairwise controlled entanglement between adjacent qubits, which propagates to create multipartite entanglement across the qubit register. This layered entanglement structure enables the quantum neural network to capture complex correlations between input features effectively. The strongly entangling layer is expressed as:

$$U_{entangle}(\theta) = \prod_{l=1}^{L}\left[\prod_{i=1}^{d} R_Z(\theta_{l,i}) CNOT(i, (i+1) \bmod d)\right] \tag{14}$$

where $R_Z(\theta_{l,i}) = e^{-i\theta_{l,i}Z/2}$ is a rotation about the Pauli-Z axis on qubit i in layer l, being a trainable parameter optimized during training, and $\mathrm{CNOT}_{i,j}$ denotes a controlled-NOT gate with control qubit i and target qubit j. The modulo operation ensures a ring topology for entanglement across all d qubits.

Quantum Measurement and Output

After transformation through the parameterized quantum gates, the measurement is performed on the computational basis. The expectation values of the Pauli-Z operators are used to obtain output features as shown below:

$$\hat{y} = \left[\langle 0|U^\dagger Z_i U|0\rangle\right]_{i=1}^{C}, \tag{15}$$

where Z_i is the Pauli-Z operator acting on the i-th qubit and C corresponds to the number of classes in the classification task. The measurement results form a vector of expectation values used as features for classification.

Cost Functions for Classification

The loss function guides the training of quantum parameters. For binary classification, we employ the binary cross entropy

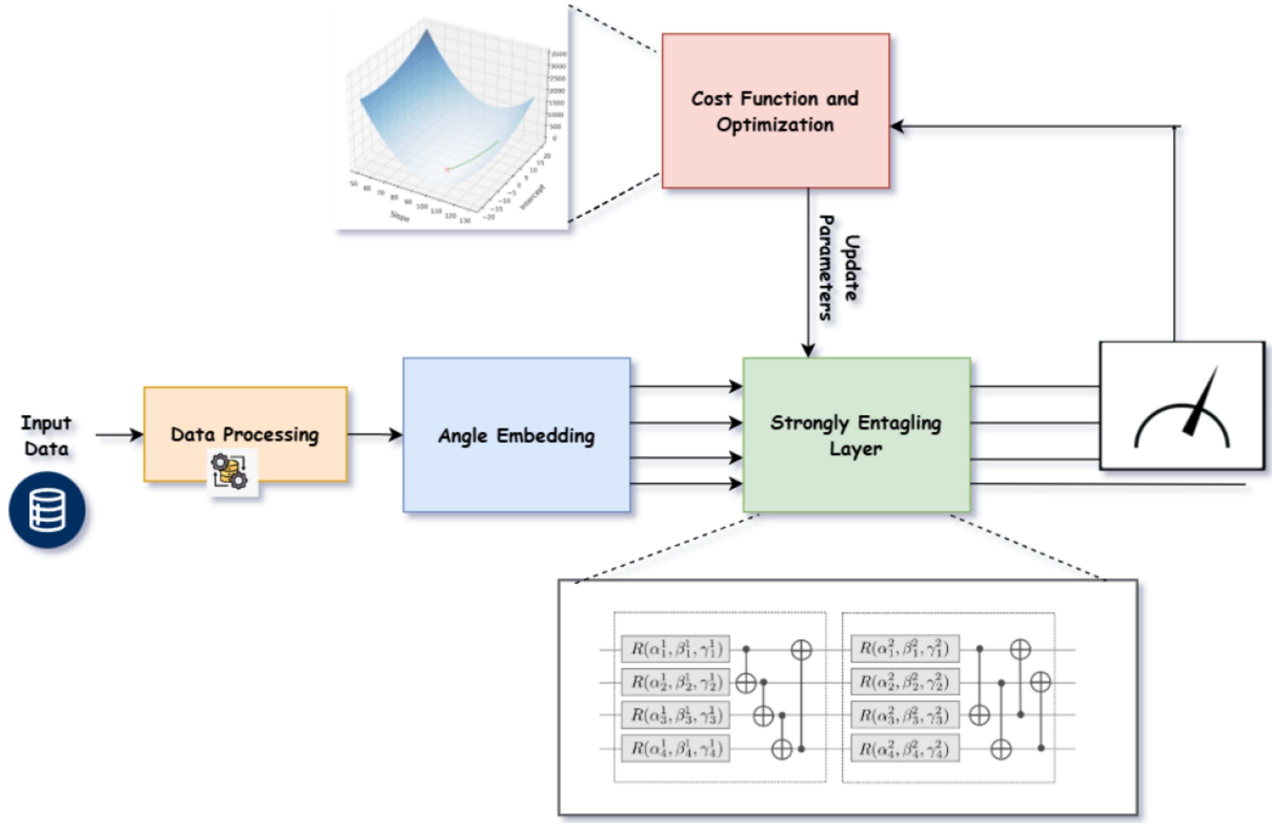


FIGURE 4: The QNN Architecture with Parameterized Quantum Circuits.

(BCE), which is given by:

$$L_{binary} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log \hat{y}_i + (1 - y_i)\log(1 - \hat{y}_i)] \quad (16)$$

where N = total number of samples, $y_i \in \{0, 1\}$ is the true label for sample i, $\hat{y}_i \in [0, 1]$ is the predicted probability for the positive class.

For multiclass classification, we utilize the categorical cross entropy (CCE) is given by:

$$L_{multi} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{i,j} \log \hat{y}_{i,j} \quad (17)$$

where C represents the total number of classes, $y_{i,j} \in \{0, 1\}$ is true class j of sample i and $\hat{y}_{i,j} \in [0, 1]$ is the predicted probability for class j of sample i. The $\hat{y}_j$ is obtained via the softmax activation function represented by:

$$\hat{y}_j = \frac{e^{z_j}}{\sum_{k=1}^{C} e^{z_k}} \quad (18)$$

This formulation allows efficient training of QNN parameters using gradient-based optimization techniques, enabling effective quantum-based classification of complex datasets.

H. NOISE MODELING AND ERROR MITIGATION IN QUANTUM K-NEAREST NEIGHBORS

QC is inherently susceptible to noise due to decoherence and gate imperfections, significantly affecting the reliability of QML algorithms such as the QKNN classifier [6] [32]. In this section, we present a noise model based on Pauli error channels and introduce an error mitigation strategy using repetition codes to improve computational stability [44] [45].

Quantum Noise Channels

Quantum noise channels describe the probabilistic evolution of the quantum state ρ under noisy conditions, which affects the precision of quantum computations [32] [46]. The primary sources of noise considered in this study include:

- **Bit Flip Channel (Pauli-X Error):** This channel applies the Pauli-X operator with probability p to the density matrix ρ, modeling bit flip errors as [6]:

$$\mathcal{E}_X(\rho) = (1 - p)\rho + pX\rho X^\dagger \quad (19)$$

- **Phase Flip Channel (Pauli-Z Error):** This channel induces phase errors, applying the Pauli-Z operator with probability p [46]:

$$\mathcal{E}_Z(\rho) = (1 - p)\rho + pZ\rho Z^\dagger \quad (20)$$

- **Bit-Phase Flip Channel (Pauli-Y Error):** This channel simultaneously introduces bit flip and phase flip errors using the Pauli-Y gate [32]:

$$\mathcal{E}_Y(\rho) = (1-p)\rho + pY\rho Y^\dagger \tag{21}$$

To analyze the robustness of QKNN under noisy conditions, we implement a probabilistic noise model where each qubit in the quantum circuit undergoes one of these errors with probability $\frac{p}{3}$ [6] [44]. The resulting noisy state evolution is given by:

$$\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Z\rho Z + Y\rho Y) \tag{22}$$

Figure 5 shows the quantum circuits for various noise models with increasing probability, ranging from $p = 0.1$ to $p = 0.6$.

Error Mitigation via Repetition Codes

To counteract the effects of quantum noise, we employ a n-qubit repetition code, an error mitigation technique that redundantly encodes logical qubits across multiple physical qubits [45].

a: Encoding and Error Detection:

A logical qubit is encoded as:

$$|0\rangle_L = |00\ldots0\rangle, \quad |1\rangle_L = |11\ldots1\rangle. \tag{23}$$

where $|0\rangle_L$ and $|1\rangle_L$ denote the logical qubit states. Errors are detected using stabilizer measurements of the form [32]:

$$S_i = Z_i Z_{i+1}, \quad i \in \{1, 2, \ldots, n-1\}. \tag{24}$$

where Z_i represents the Pauli-Z operator on the i-th qubit; measurement outcomes $+1$ indicate no error and -1 signal bit flip or phase flip errors. Table 5 shows the error syndrome table for a 3-qubit repetition code, illustrating the possible error patterns and their corresponding corrections.

TABLE 5: Error Syndrome Table for a 3-Qubit Repetition Code.

S_1	S_2	Error Pattern	Correction
+1	+1	$I \otimes I \otimes I$	No Error
+1	-1	$I \otimes I \otimes X$	Apply X gate on Qubit 1
-1	+1	$X \otimes I \otimes I$	Apply X gate on Qubit 3
-1	-1	$I \otimes X \otimes I$	Apply X gate on Qubit 2

b: Error Correction:

By analyzing stabilizer syndromes, errors are identified and corrected using a majority voting scheme [47]:

- If the majority of qubits remain in $|0\rangle$, the logical state is $|0\rangle_L$.
- If the majority of qubits are in $|1\rangle$, the logical state is $|1\rangle_L$.

This is a redundancy-based approach, which ensures resilience to single qubit errors and improves the reliability of QKNN classification under noisy conditions [6]. Although repetition coding inherently increases the number of quantum gates and thus the overall circuit depth, potentially introducing additional noise, the technique leverages the statistical properties of independent error occurrences to reduce the effective logical error rate. Assuming that the physical qubit errors are independent and identically distributed with probability p, the logical error corresponds to the event where the majority of the n physical qubits encoding a single logical qubit are erroneous. This can be quantitatively expressed as:

$$P_{\text{logical error rate}} = \sum_{k=\lceil \frac{n}{2} \rceil}^{n} \binom{n}{k} p^k (1-p)^{n-k} \tag{25}$$

where $\lceil \frac{n}{2} \rceil$ is the smallest integer greater than or equal to half of n. Since this is the tail probability of a binomial distribution, $P_{\text{logical error rate}}$ decreases exponentially with increasing n, provided $p < 0.5$. Thus, despite the increased gate count and potential noise sources from additional operations, the net effect of repetition coding is an exponential suppression of logical errors relative to physical errors. This statistical advantage justifies the use of repetition codes as an effective error mitigation strategy, though it is essential to balance the trade-off between improved noise resilience and the overhead of deeper circuits, especially in the context of current NISQ devices.

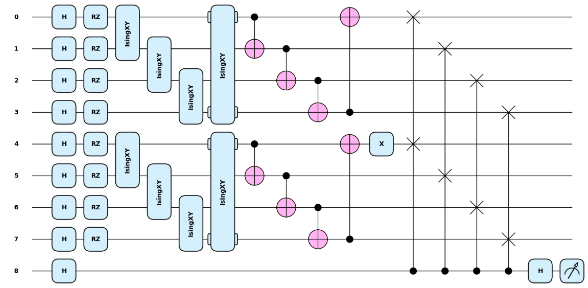
The integration of repetition codes into QKNN enables error-resilient quantum state preparation and measurement, which is essential for practical implementations on near-term quantum devices. Future research may explore more advanced error correction strategies, such as surface codes, to further enhance the robustness of QML algorithms [47].

IV. RESULT AND DISCUSSION

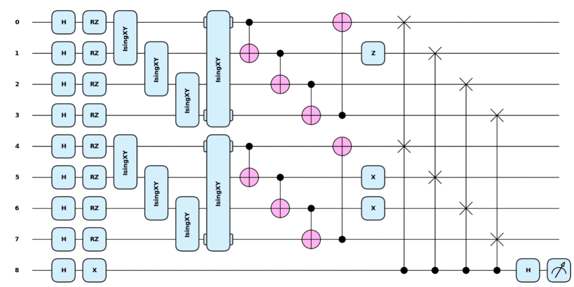
In this part, we conduct a thorough evaluation of the classification capabilities of QKNN, CKNN, and QNN on the three standard datasets: Breast Cancer, Iris, and Banknote Authentication. The review provides a comprehensive picture of the success of each model by including important performance parameters, including accuracy, precision, recall, F1 score, and AUC. Furthermore, in order to evaluate the stability of the model, we investigate how quantum noise affects the performance of QKNN.

A. COMPARATIVE PERFORMANCE ANALYSIS ACROSS MODELS

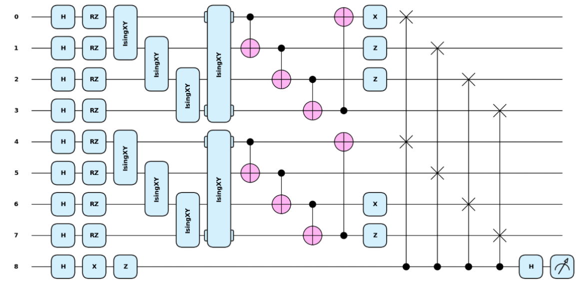
A comparative performance of the proposed QKNN, QNN and CKNN classification accuracies is presented in Table 6 for the three test datasets. The proposed QKNN consistently outperforms or matches CKNN, showcasing its superior classification capabilities in diverse datasets. While the QNN model exhibits competitive performance in several datasets, its overall accuracy is somewhat limited compared to QKNN. This discrepancy arises primarily from the inherent challenges in optimizing variational quantum circuits, which form the core of QNN architectures. Such optimization difficulties



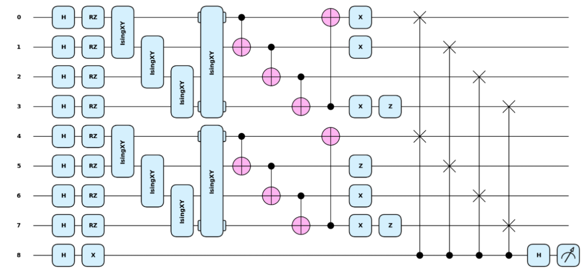
(a) Noise probability $p = 0.1$



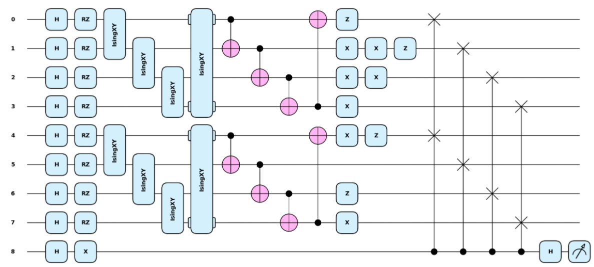
(b) Noise probability $p = 0.2$



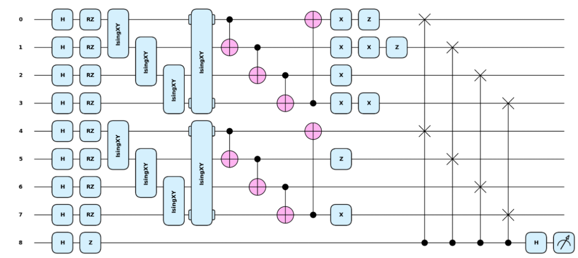
(c) Noise probability $p = 0.3$



(d) Noise probability $p = 0.4$



(e) Noise probability $p = 0.5$



(f) Noise probability $p = 0.6$

FIGURE 5: The quantum circuits for various noise models with increasing probability.

are especially pronounced in scenarios with limited training data, where the model struggles to converge to optimal parameter settings. Consequently, QNN may not fully realize its potential in these contexts, resulting in comparatively lower classification accuracy.

TABLE 6: Classification Accuracies of the Proposed QKNN, QNN and CKNN Algorithms for the Three Test Datasets.

Dataset	QKNN	CKNN	QNN
Breast Cancer	**0.9825**	0.9298	0.9717
Iris	**1.0000**	**1.0000**	0.8333
Bank Note	**0.9927**	0.9855	0.8618

The results demonstrate that the proposed QKNN algorithm consistently outperforms CKNN, particularly on the Wisconsin Breast Cancer dataset. Specifically, QKNN achieves an accuracy of 98.25%, surpassing both CKNN (92.98%) and QNN (97.17%). This superior performance is attributed to the utilization of entangling gates such as IsingXY and CNOT, which facilitate richer quantum state representations and enhance the model's capacity to discriminate complex and overlapping class boundaries. In particular, QNN also outperforms CKNN in this dataset, suggesting that variational quantum circuits can effectively capture the salient features of the data despite the inherent challenges of optimization.

For the Iris dataset, both QKNN and CKNN achieve a perfect classification accuracy of 100%, indicating the inherent linear separability of the dataset. In contrast, QNN performs significantly worse, attaining an accuracy of only 83.33%. This disparity suggests that QNN struggles to effectively capture the underlying feature space, likely due to the limitations of quantum variational circuits when applied to small-scale training scenarios.

Similarly, in the Bank Note Authentication dataset, QKNN achieves the highest accuracy of 99.27%, surpassing both

CKNN (98.55%) and QNN (86.18%). The lower performance of QNN in this case further underscores the difficulty in training QNNs compared to distance-based quantum-enhanced methods such as QKNN. For better visualization, the comparative performance analysis included in Table 6 is also shown in the bar diagram of Figure 6.

B. KEY PERFORMANCE METRICS AND THEIR IMPLICATIONS

To further assess the efficiency of the proposed QKNN algorithm, we analyze additional performance metrics, such as AUC, F1 score, precision, and recall, which are shown in Table 7. These metrics provide deeper insight into the reliability of the classification beyond accuracy.

TABLE 7: Additional Performance Metrics of the Proposed QKNN Algorithm.

Dataset	AUC	F1-Score	Precision	Recall
Breast Cancer	0.9859	0.98	0.98	0.99
Iris	1.0000	1.00	1.00	1.00
Bank Note	0.9921	0.99	0.99	0.99

The consistently high AUC values and balanced F1 scores demonstrate that QKNN possesses strong discriminative power and effectively balances precision and recall across all datasets, minimizing both false positives and false negatives. This balance is critical in high-stakes applications such as medical diagnosis and fraud detection, where misclassification entails severe consequences. These superior results stem from the quantum-enhanced feature encoding and distance computation strategies employed by QKNN, which yield richer data representations and more precise similarity assessments than classical approaches. In particular, the use of entangling gates improves class separability within the quantum feature space. Taken together, these performance metrics affirm the robustness and reliability of QKNN in noisy and complex data environments.

C. CONFUSION MATRICES FOR MODEL PERFORMANCE

Confusion matrices provide additional insights into the ability of a model to distinguish between true positives, false positives, true negatives, and false negatives. The confusion matrices for the three dataset are shown in Figure 7. The visualizations enable a granular understanding of the performance of the model beyond the aggregate metrics.

The confusion matrices show that QKNN consistently achieves high true positive and true negative rates while minimizing false positives and false negatives, particularly on complex datasets such as Breast Cancer and Banknote Authentication. This confirms the strong discriminative capacity of the model and its reliable and balanced classification performance, highlighting its potential for practical high-stakes quantum machine learning applications.

D. BENCHMARKING AGAINST EXISTING QUANTUM KNN APPROACHES

We benchmark the proposed Quantum k-Nearest Neighbors (QKNN) algorithm against recent state-of-the-art quantum KNN methods using three widely recognized datasets: Wisconsin Breast Cancer, Iris, and Bank Note Authentication. However, during benchmarking, we found a limited availability of comparable studies on these specific datasets, which constrained direct performance comparisons with existing literature.

The proposed QKNN algorithm achieves classification accuracies of 98.25%, 100.00% and 99.27% for the Wisconsin Breast Cancer, Iris and Bank Note Authentication datasets, respectively. The high performance has been achieved by enhancing the feature encoding using the superposition and entanglement phenomena and by exploiting the distance computation. For performance comparison, Table 8 shows the classification accuracies of the proposed QKNN algorithm along with those reported in recent QKNN studies by Feng et al. [7], Maldonado et al. [8], Li et al. [9] and Bhaskaran et al. [10].

TABLE 8: Performance Comparison of the Proposed QKNN Algorithm with Existing Studies

Reference	Wisconsin Breast Cancer	Iris	Bank Note Authentication
Feng et al. [7]	–	95.82%	–
Meldonado et al. [8]	–	94%	87.50%
Li et al. [9]	–	–	–
Bhaskaran et al. [10]	93.85%	93.33%	–
Proposed QKNN	**98.25%**	**100%**	**99.27%**

The performance of the proposed QKNN algorithm exceeds that of the existing QKNN algorithms for all three test datasets. The former improves the accuracy by 4.40% over the QKNN algorithm by Bhaskaran et al. for the Wisconsin Breast Cancer dataset and achieves the perfect classification accuracy for the Iris dataset, outperforming all previously reported quantum benchmarks. Furthermore, in the Bank Note Authentication dataset, our approach establishes a new state-of-the-art quantum performance, improving the performance of Maldonado et al. by 11.77%. These results highlight the effectiveness of our quantum-enhanced feature encoding and noise-resilient distance metric within the quantum computational framework. This comprehensive evaluation demonstrates the significant advancement that the proposed QKNN algorithm offers in QML classification.

E. EFFECT OF QUANTUM NOISE ON QKNN PERFORMANCE

QC systems are susceptible to noise, which can have a negative impact on the performance of classification. Figure 8 illustrates the effect of increasing the noise probability on the accuracy of the proposed QKNN algorithm.

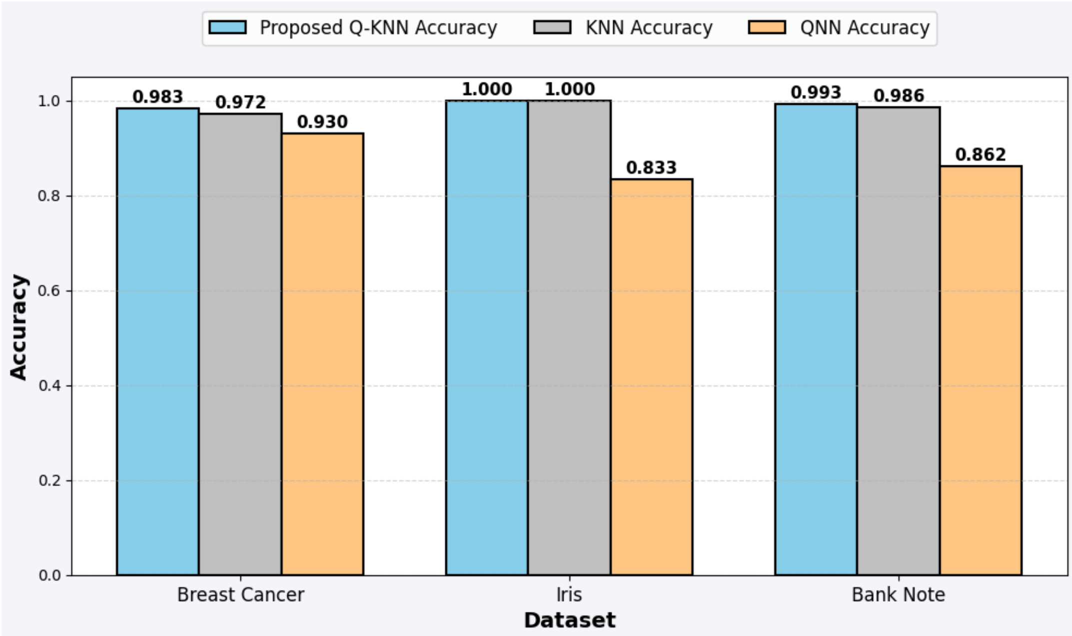


FIGURE 6: Comparative Performance of QKNN, CKNN, and QNN across the three datasets.

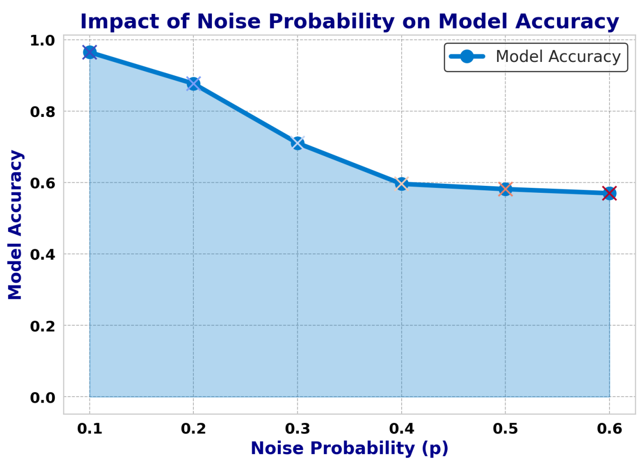


FIGURE 8: Impact of Noise Probability on Model Accuracy.

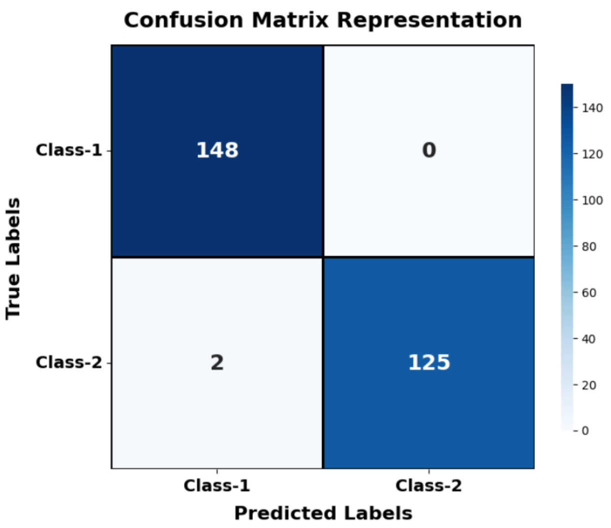
The noise probability p is varied from 0 to 0.6 to observe its impact on the performance of the model. The results highlight a clear inverse relationship between noise probability and model accuracy. QKNN achieves an accuracy of 96.49% at a low noise level of $p = 0.1$, demonstrating its robustness under minimal noise. However, as p increases to 0.3, accuracy declines sharply to 71.05%, indicating the sensitivity of the model to quantum decoherence and gate errors. At higher noise levels ($p \geq 0.6$), accuracy stabilizes at approximately 57%, suggesting a noise tolerance threshold beyond which model performance is severely degraded.

These observations emphasize the importance of implementing quantum error correction techniques such as repetition encoding to mitigate the effects of noise and improve the robustness of classification. The need for error correction becomes more evident as the noise probability increases, underscoring the challenges faced by QML models in noisy quantum environments.

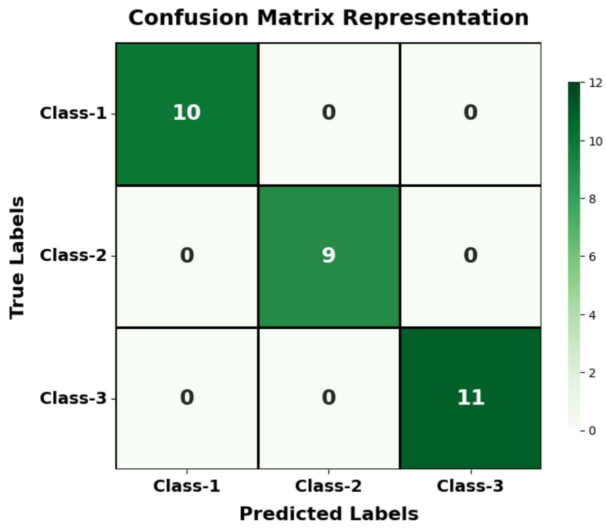
V. CONCLUSION

This study presents a novel QKNN algorithm as a quantum-enhanced classification method and evaluates its performance against CKNN and QNN in three test datasets, namely Wisconsin Breast Cancer, Iris, and Bank Note Authentication. The experimental results show that the proposed QKNN algorithm outperforms or matches the CKNN, particularly excelling in datasets with complex decision boundaries. Specifically, QKNN achieves an accuracy of 98.25% in the breast cancer dataset compared to 92.98% in CKNN and achieves 99.27% accuracy in the banknote dataset, outperforming CKNN at 98.55%. The results demonstrate the ability of QKNN to better handle complex and high-dimensional classification tasks. Again, for the Iris dataset, both QKNN and CKNN achieved a perfect classification accuracy of 100%, while QNN lagged behind at 83.33%, suggesting that variational quantum circuits may not be very effective for certain classification tasks without extensive parameter optimization. Furthermore, benchmarking against recent QKNN studies shows that our approach improves accuracy by up to 11.77% over the former. Thus, the proposed QKNN may prove to be a new state of the art technique for the quantum K-nearest neighbor classification method in the standard datasets.

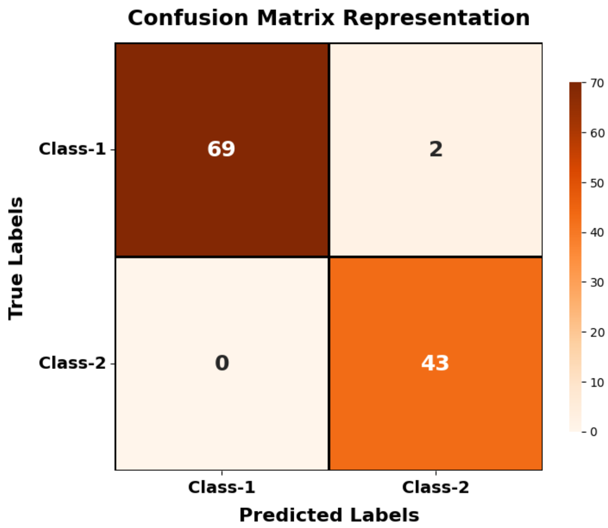
A critical aspect of this study was to analyze the consequences of quantum noise on QKNN accuracy. It should be noted here that the results reveal a strong inverse correlation between noise probability and model accuracy, with performance sharply declining beyond $p = 0.3$, stabilizing



(a) Confusion Matrix for Banknote Authentication Dataset



(b) Confusion Matrix for Iris Dataset



(c) Confusion Matrix for Breast Cancer Wisconsin Dataset

FIGURE 7: Confusion Matrices for the Three Datasets.

at approximately 57% accuracy at $p = 0.6$. This highlights the sensitivity of quantum models to noise and the hurdles posed by quantum decoherence and gate errors on near-term quantum devices. These findings further emphasize the need for incorporating quantum error correction strategies to improve the performance of the model in practical quantum environments.

The implications of these results are significant for QML applications. The superior accuracy and robustness of QKNN in complex datasets suggest its potential in critical areas such as medical diagnostics, financial fraud detection, and security. However, to move from theoretical promise to practical deployment, several avenues need to be explored. Firstly, optimizing quantum feature encoding techniques to improve scalability on larger datasets is essential. Secondly, developing adaptive quantum circuits that dynamically adjust to noise conditions could further enhance QKNN's performance. Additionally, empirical validation on real quantum hardware will provide valuable insights into understanding quantum computational constraints, guiding future refinements of the model. Finally, integrating QKNN into hybrid quantum-classical frameworks will allow the strengths of both paradigms to be leveraged, potentially leading to more powerful and scalable QML solutions. As QC technology matures, continued improvements in QKNN and related quantum classifiers may revolutionize data driven decision-making in high impact industrial applications.

VI. ACKNOWLEDGMENT

The authors express their sincere gratitude to the Department of Electrical and Electronic Engineering (EEE) at the Bangladesh University of Engineering and Technology

(BUET) for providing the necessary resources, support and a conducive environment to conduct this research.

REFERENCES

[1] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, 2017.

[2] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[3] N. Wiebe, A. Kapoor, and K. M. Svore, "Quantum algorithms for nearest-neighbor methods for supervised and unsupervised learning," *Quantum Information and Computation*, vol. 15, no. 3-4, pp. 316–356, 2015.

[4] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, no. 7747, pp. 209–212, 2019.

[5] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum algorithms for supervised and unsupervised machine learning," *arXiv preprint arXiv:1307.0411*, 2013.

[6] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[7] C. Feng, B. Zhao, X. Zhou, X. Ding, and Z. Shan, "An enhanced quantum k-nearest neighbor classification algorithm based on polar distance," *Entropy*, vol. 25, no. 1, p. 127, 2023.

[8] A. Maldonado-Romo, J. Y. Montiel-Pérez, V. Onofre, J. Maldonado-Romo, and J. H. Sossa-Azuela, "Quantum K-Nearest Neighbors: Utilizing QRAM and SWAP-Test Techniques for Enhanced Performance," *Mathematics*, vol. 12, no. 12, p. 1872, 2024.

[9] J. Li, J. Zhang, J. Zhang, and S. Zhang, "Quantum knn classification with k value selection and neighbor selection," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 43, no. 5, pp. 1332–1345, 2023.

[10] P. Bhaskaran and S. Prasanna, "An accuracy analysis of classical and quantum-enhanced k-nearest neighbor algorithm using canberra distance metric," *Knowledge and Information Systems*, vol. 67, no. 1, pp. 767–788, 2025. [Online]. Available: https://doi.org/10.1007/s10115-024-02229-w

[11] Z. Yang, M. Zolanvari, and R. Jain, "A survey of important issues in quantum computing and communications," *IEEE Communications Surveys & Tutorials*, vol. PP, no. 99, pp. 1–1, Jan. 2023.

[12] R. D. Ranga, A. Rana, S. Prajapat, P. Kumar, K. Kumar, and A. V. Vasilakos, "Quantum machine learning: Exploring the role of data encoding techniques, challenges, and future directions," *Mathematics*, vol. 12, no. 21, p. 3318, 2024.

[13] H. Yano, Y. Suzuki, K. M. Itoh, R. Raymond, and N. Yamamoto, "Efficient discrete feature encoding for variational quantum classifier," *arXiv preprint arXiv:2005.14382*, 2020.

[14] X. B. Nguyen, H. Q. Nguyen, H. Churchill, S. U. Khan, and K. Luu, "Quantum visual feature encoding revisited," *arXiv preprint arXiv:2405.19725*, 2024.

[15] K. Sakka, K. Mitarai, and K. Fujii, "Automating quantum feature map design via large language models," *arXiv preprint arXiv:2504.07396*, 2025.

[16] Z. Jarir and M. Quafafou, "Quantum-enhanced k-nearest neighbors for text classification: A hybrid approach," *Advanced Quantum Technologies*, vol. 2024.

[17] E. Zardini, E. Blanzieri, and D. Pastorello, "A quantum k-nearest neighbors algorithm based on the euclidean distance estimation," *Quantum Machine Intelligence*, vol. 2024.

[18] A. Berti, A. Bernasconi, G. M. Del Corso, and R. Guidotti, "The role of encodings and distance metrics for the quantum nearest neighbor," *Quantum Machine Intelligence*, vol. 6, no. 2, p. 62, 2024.

[19] Y. Wang and R. Wang, "Improved handwritten digit recognition using quantum k-nearest neighbor algorithm," *Springer*, 2019.

[20] P. Bhaskaran and S. Prasanna, "An accuracy analysis of classical and quantum-enhanced k-nearest neighbor algorithm," *Knowledge and Information Systems*, 2025.

[21] Q. Xiang, D. Li, Z. Hu, Y. Yuan, Y. Sun, Y. Zhu, Y. Fu, Y. Jiang, and X. Hua, "Quantum classical hybrid convolutional neural networks for breast cancer diagnosis," *Scientific Reports*, vol. 14, no. 1, p. 24699, 2024.

[22] R. Zhul, G. Li, J.-X. Liu, L.-Y. Dai, S. Yuan, and Y. Guo, "A fast quantum clustering approach for cancer gene clustering," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2018, pp. 1610–1613.

[23] W. Street, W. Wolberg, and O. Mangasarian, "Nuclear feature extraction for breast tumor diagnosis," UCI Machine Learning Repository, 1993.

[24] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

[25] D. Dua and C. Graff, "Banknote authentication dataset," UCI Machine Learning Repository, 2019.

[26] G. Forman, "An extensive empirical study of feature selection metrics for text classification," *Journal of Machine Learning Research*, vol. 3, pp. 1289–1305, 2003.

[27] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, 2018.

[28] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, "Circuit-centric quantum classifiers," *Physical Review A*, vol. 101, no. 3, Mar. 2020. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.101.032308

[29] P.-L. Dallaire-Demers and N. Killoran, "Quantum generative adversarial networks," *Physical Review A*, vol. 98, no. 1, p. 012324, 2018.

[30] S. Aaronson, *Read the fine print*, 2015, vol. 11, no. 4.

[31] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," 2008.

[32] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge university press, 2010.

[33] D. Deutsch, "Quantum theory, the church-turing principle, and the universal quantum computer," *Proceedings of the Royal Society of London A*, vol. 400, no. 1818, pp. 97–117, 1989.

[34] R. P. Feynman, "Quantum mechanical computers," *Foundations of Physics*, vol. 16, no. 6, pp. 507–531, 1986.

[35] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484–1509, 1997.

[36] L. K. Grover, "A fast quantum mechanical algorithm for database search," *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pp. 212–219, 1996.

[37] L.-Z. Gao, C.-Y. Lu, G.-D. Guo, X. Zhang, and S. Lin, "Quantum k-nearest neighbors classification algorithm based on mahalanobis distance," *Frontiers in Physics*, vol. 10, p. 1047466, 2022.

[38] J. Li, S. Lin, K. Yu, and G. Guo, "Quantum k-nearest neighbor classification algorithm based on hamming distance," *Quantum Information Processing*, vol. 21, no. 18, 2022.

[39] M. Schuld, I. Sinayskiy, and F. Petruccione, "The quest for a quantum neural network," *Quantum Information Processing*, vol. 14, no. 7-8, pp. 796–817, 2014.

[40] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Physical Review A*, vol. 98, no. 3, p. 032309, 2018.

[41] R. Zhao and S. Wang, "A review of quantum neural networks: methods, models, dilemma," *arXiv preprint arXiv:2109.01840*, 2021.

[42] M. Schuld and F. Petruccione, *Machine learning with quantum computers*. Springer, 2021, vol. 676.

[43] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational quantum algorithms," *Nature Reviews Physics*, vol. 3, no. 9, p. 625–644, Aug. 2021. [Online]. Available: http://dx.doi.org/10.1038/s42254-021-00348-9

[44] S. J. Devitt, W. J. Munro, and K. Nemoto, "Quantum error correction for beginners," *Reports on Progress in Physics*, vol. 76, no. 7, p. 076001, 2013.

[45] D. A. Lidar and T. A. Brun, *Quantum error correction*. Cambridge university press, 2013.

[46] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*. Oxford University Press, 2002.

[47] D. Gottesman, "Stabilizer codes and quantum error correction," Ph.D. dissertation, California Institute of Technology, 1997.

ASIF AKHTAB RONGGON is currently pursuing a Bachelor of Science (B.Sc.) degree in the field of Electrical and Electronics Engineering from Bangladesh University of Engineering and Technology (BUET).

His academic background in quantum computing, quantum communication, machine learning, and signal processing is solid. His research interests cover quantum machine learning, quantum communication, quantum information processing, and quantum error correction.

MD. SAIFUR RAHMAN received his B.Sc. and M.Sc. degrees in Electrical and Electronic Engineering from BUET in 1987 and 1989, respectively. He earned his Ph.D. in Digital Signal Processing from UMIST (now the University of Manchester), UK, in 1992 as a Commonwealth Scholar. He also completed a postdoctoral fellowship in renewable energy and sustainable technologies at Curtin University, Australia (1999-2001).

He joined BUET as a faculty member in 1987 and has held various academic and administrative positions, including Head of the Department of EEE and Dean of the Faculty of EEE. His research interests encompass digital communication, signal processing, quantum communication, intelligent networks, embedded systems, and design of renewable energy systems. He is a member of various national and international institutes and societies. He has published extensively in international journals and conferences and has contributed to consulting and sponsored research projects in the telecommunications and energy sectors.