IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 16 November 2022, accepted 12 December 2022, date of publication 26 December 2022,
date of current version 30 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3232307

RESEARCH ARTICLE

A Fully Connected Quantum Convolutional Neural Network for Classifying Ischemic Cardiopathy

UBAID ULLAH[1], ALAIN GARCíA OLEA JURADO[2], IGNACIO DIEZ GONZALEZ[2], AND BEGONYA GARCIA-ZAPIRAIN[1], (Member, IEEE)

[1]eVIDA Research Group, University of Deusto, 48007 Bilbao, Spain
[2]Cardiology Service, Basurto University Hospital, Bilbao, 48013 Bizkaia, Spain

Corresponding author: Ubaid Ullah (ubaid.ullah@deusto.es)

This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie grant agreement Number 847624. In addition, a number of institutions back and co-finance this project. Furthermore, the Basque government to the eVIDA research group IT 1536-22, University of Deusto and coordinated by Basurto Hospital in Bilbao, which provides the database. The Clinical Research Ethics Committee of Euskadi (PI202031) in Spain validated the research protocols. The article reflects only the author's view, and the Agency is not responsible for any use that may be made of the information it contains.

ABSTRACT The prevalence of heart diseases is rising quickly throughout the world, which has an impact on both the world economy and public health. According to the recent statistical survey reports, the increasing mortality rate is due to high blood pressure, high cholesterol, the use of tobacco, obesity, and an inconsistent pulse rate. It is difficult and time-consuming to investigate the various variations of these factors and their impact on Coronary Artery Disease (CAD). Therefore, it is necessary to use modern approaches to diagnose the disease early and minimize the mortality rate. The fields of machine learning and data mining have a wide research dimension and various novel techniques that could help in the prediction of CAD in its early stages and identify their patterns and behaviors in a huge amount of data. The results of such predictions will aid the clinical staff in decision making and early diagnosis. In such a scenario, we proposed a quantum version of the Fully Convolutional Neural Network (FCQ-CNN) for Ischemic Heart Disease (IHD) classification. The proposed model evaluates the quantum circuit-based technique that was inspired by convolutional neural networks, a very successful machine learning model. This method provides $O(log(n))$ depth for n qubits, reducing the number of parameters and allowing for effective training and testing of real quantum devices. The model has been evaluated by considering the IHD dataset after the data has been cleaned and filtered through the Maximally Relevant Minimally Redundancy (MRMR) filter. For dimension reduction, a Support Vector Machine along with Recursive Feature Elimination (SVM-RFE) has been considered. Initially, the model is tested with 20% of the whole dataset and gets the promising results of a testing accuracy of 84.6% with a testing loss of 0.28. By taking into account the same optimal parameters, the proposed model outcomes are compared to those of the classical Optimized Convolution Neural Network (Optimized-CNN) and Fully Connected Neural Network (FCNN) models. Comparing the model's competency to that of earlier published quantum models yields improvements in accuracy of 8.6%, 12.6%, 3.5%, and 1.8% respectively.

INDEX TERMS Ischemic cardiopathy classification, feature selection, quantum machine learning algorithm, hyperparameters optimization.

I. INTRODUCTION

Ischemic heart disease has become a major health concern for many patients due to its high mortality rate throughout the world [1]. Early diagnosis of cardiac diseases has the potential to save many lives, whereas routine clinical data analysis presents significant challenges in identifying IHD disorders such as heart attacks, strokes, heart failure, etc [2], [3]. The quick rise of coronary heart disease is more common in males than females, especially in middle aged or older persons which poses a serious risk to public health and places a heavy burden on global healthcare systems [4], [5]. Additionally, the children in developed countries have an increasing CAD risk [6]. According to the World Health

The associate editor coordinating the review of this manuscript and approving it for publication was M. Shamim Kaiser.

Organization (WHO), every year globally one third of the population, or around 17.9 million people, die from IHD, with Asia having the highest prevalence [7], [8]. The European Cardiology Society (ESC) reported that each year, 3.6 million adults globally suffer from cardiac disease, and 26 million have been diagnosed with coronary heart diseases [9]. The various unhealthy behaviors, such as diabetes, rising cholesterol, obesity, hypertension, an increase in triglycerides, and others, increase the risk of developing coronary heart disease [7]. Moreover, there are also other symptoms that the American Heart Association lists, such as difficulty sleeping, an irregular heartbeat, difficulties breathing, swollen legs, and occasionally weight gain that happens rapidly, up to 1-2 kg each day [10], [11]. All of these symptoms are indicative of various illnesses, including those that affect elderly people, and are more challenging to diagnose at an early stage.

Currently, echocardiographic stress tests and pharmacological stress tests are used to diagnose IHD by performing a functional evaluation of the coronary artery. A coronary angiography, which provides information on the anatomical structure of the heart, is another diagnostic imaging test for CAD. However, because of the complexity of the coronary anatomy and the structure of the plaque, there is still a chance that the disease will be erroneously diagnosed [12]. Additionally, patients may encounter the possibility of health issues with the use of general anesthesia and contrast material. Estimating the risk of cardiac failure accurately aids in patient protection and the avoidance of severe cardiac strokes [13]. Researchers now have access to massive dataset to create the most precise prediction model thanks to the development of artificial intelligence and machine learning. When trained on relevant data, machine learning algorithms are capable of accurately diagnosing diseases [14]. Recent research, which focuses on cardiac disease in both children and adults, highlights the importance of decreasing mortality from IHD. As a result, it is necessary to investigate the effects of risk factors, their high prevalence in the general population, their significant impact on individual cardiac diseases, and their treatment capability to reduce cardio diseases [15].

Furthermore, the selection of the best features is an essential step to keep in mind when applying machine learning (ML) techniques for the prediction of a certain disease. Initially, the input dataset in their current form are inconsistent and contain missing data, duplicate entries, and other redundant information [16]. In model prediction, it is important to choose the strongest and most important attributes that may be used for the identification or classification of the targeted disease. For the development of accurate prediction, effort should be made to choose the important feature combinations, which help make the machine learning process more accurate and maximize the prediction power of the model [17]. During modeling the predictors for such diseases, different studies consider different risk factors or features as input for their models depending on their feature selection techniques [18], [19]. When implementing the IHD prediction models, various factors are taken into consideration, i.e., high cholesterol, age, chest pain, fasting blood pressure, diabetes, results from a resting electrocardiogram, exercise induced angina, depression, heart status, maximum heart rate, family history, poor diet, obesity, alcohol consumption, and inactivity [20], [21]. These factors play an important role in disease classification and detection.

Recently, numerous quantum enhanced machine learning algorithms have been introduced to speed up some specific ML tasks [22], [23]. The use of quantum deep learning (QDL) and ML models for medical applications has gained significant attention in the past few years [24], [25]. In addition, various researchers have previously used different quantum machine learning algorithms [26], [27] by utilizing real-time healthcare datasets [28], [29]. The academic community and the medical field have both recently shown a substantial interest in supervised learning [30]. The quantum supervised machine learning algorithms, i.e, Variational Quantum Classifier (VQC) [31], [32], Quantum Support Vector Machine (QSVM) [33], the robust pre-processing techniques [34], [35] and error reducing algorithms [36], [37] has a numerous contributions in coping with the classification problems. Convolutional neural network models (CNNs) are considered the most successful machine learning algorithms for pattern recognition and disease prediction. Furthermore, CNNs have shown enhanced performance in early disease diagnosis, which is rooted in the concept of managing the input dataset through numerous representational layers [38], [39]. Many hybrid algorithms use parameterized circuit optimization to reduce the cost function, which usually consists of the expectation value of an output observable, which is the same as classical model parameter optimization [40], [41], [42]. From the other side, by using the concept of quantum supremacy, quantum computing (QC) has demonstrated its importance for solving problems. The frequent development of QC in the field of ML can naturally handle several problems with complicated input correlations that can be exceptionally difficult for classical computers [43], [44], [45]. Moreover, near-term quantum computers have also shown a tremendous impact from quantum computing on ML [46].

In this research, we aim to construct a fully connected quantum convolutional neural network model, demonstrating the application of various methods for early stage diagnosis of IHD. The proposed model uses a quantum circuit based technique, which reduces the model's complexity by reducing the number veriational parameters. Multiple qubit gates are applied to adjacent qubits in the convolution circuit to find the hidden state, and the pooling circuit reduces the model complexity by considering the CNOT gate operation. The dataset used for this work is private, which is provided by a hospital contains on health record of patients. We attempted to classify the desired disease by using the labels obtained from the input dataset. At first, we tried to explain how disease classification works by utilizing the feature extraction method. Later on, to increase the chances of success, we combine the classical and quantum versions of the convolutional neural network algorithm with our input dataset, which consists of features

related to IHD, i.e., hypercholesterolemia, smoking habits, antipsychotic drugs, heart failure, and so on. Comparing the proposed quantum model to a classical Optimized-CNN and FCNN algorithm, the classification results achieved are quite promising. The primary objective of this research is to propose a robust model by combining the benefits of quantum computing with CNNs, and clinical record data is considered a key tool. This results in the least expensive model for early stage IHD diagnosis. The proposed model helps to enhance the CNN performance and ensure the accurate prediction of IHD, and it is also useful to speed up the diagnosis process for heart patients in the early stages.

II. MATERIALS

The dataset used for this research work has been provided by Basurto Hospital of the Basque Country and includes patient health record information. For patients suspected of having coronary heart disease, the pretest probabilities should be predicted based on, symptoms, gender, age, BMI, etc. The main patient level data collection tools in healthcare are Electronic Health Records (EHR), which can also be used for research-related purposes. To enhance the pretest IHD risk adjustment more precisely, we may examine which EHR features have greater connections with heart failure and cardiac problems. The gathered dataset consists of patient records of both genders collected between 2016 and 2018. This data was gathered for approximately 2,199,711 patients, among those 43,835 individuals who needed treatment for chest pain and thoracic pain within the particular time period. Initially, there were 82 attributes in the dataset, including the targeted CAD variable, and each variable was either categorical or continuous. The dataset has 42 effective features, are described below [47].

- Catheterization: The process in which a flexible tube is guided through a blood vessel to diagnose clogged arteries, irregular heartbeats, etc.
- Ergometry: Ergometry is the measurement of human physical performance.
- Echo Stress: A test that uses ultrasound imaging to show how well the heart muscle works at pumping blood to the body in a stressful situation.
- Echocardiogram: It represents the graphic outline of the heart's movement.
- ECG: The heart's electrical activity is recorded by an ECG. It is a frequent and painless test for identifying heart issues and keeping tabs on heart health.
- Depression: Depression is characterized by a constant feeling of sadness and a lack of interest. It is a medical illness that affects a person's mood and ability to function.
- Alcohol: Alcohol intake causes illness when a person's body develops a dependence on or addiction to it.
- Drug dependence: A biopsychosocial situation whereby a person's functionality is dependent on the consumption, meaning cessation of repeated exposure to a stimulus.
- Anxiety: Mental health condition which carries excessive anxiety or worries, about several things such as personal health, work, and social interactions.
- Dementia: Refers to abnormal brain alterations that affect memory, language, and problem-solving skills in addition to other aspects of thinking.
- Renal insufficiency: Renal failure is a medical condition in which the kidneys stop working less than normal levels.
- Diabetes mellitus: To put it more simply, diabetes mellitus occurs when there is an abnormally high level of sugar flowing in the bloodstream of a person.
- Diabetes type 1: A syndrome when the body is unable to create enough insulin, causing the blood glucose level to rise too high.
- Diabetes type 2: A condition when the blood glucose level is excessively high because the body cannot respond to insulin and frequently produces insufficient insulin.
- Dyslipidemia: The blood's lipid (fat) content is abnormally high.
- Hypercholesterolemia: Hypercholesterolemia is a type of dyslipidemia that is described as a high level of cholesterol in the blood.
- Ventricular fibrillation: Refers to a type of arrhythmia, or irregular heartbeat, affecting the heart's ventricles.
- Auricular Flutter: Refers to another type of arrhythmia which causes the heart to beat faster and inefficiently.
- Heart failure: A disorder when the heart is unable to adequately circulate blood throughout the body.
- Obesity: Overweight and obesity are described as accumulations of abnormal or excessive fat that pose a health risk.
- Age: Number of years a person has lived
- Antithrombotic agents: Antithrombotic agents are drugs that reduce the formation of blood clots.
- Acetylsalicylic acid: To avoid blood clots, strokes, and myocardial infarction (MI), doctors prescribe analgesic drugs that stop platelet aggregation.
- Protonpump inhibitors: Anti acid medications known as proton pump inhibitors act by lowering stomach acid production.
- Diuretics: Diuretics are drugs that work by increasing the amount of water and salt that the body excretes in the form of urine.
- Hormonal Contraceptives: A type of birth control that uses hormones such as estrogen and progesterone by blocking the release of eggs from ovaries.
- Anabolic steroids: A synthetic version of testosterone, which increases protein within cells, is used for growth muscle stimulation.
- Antigout Preparations: Agents that increase uric excretion by the kidney, decrease uric acid production.
- Immunomodulators: A group of drugs that work includes working on the immune system directly, by turning down some proteins and turning up others.

- Antidepressants: Refers to the category of medications used to treat anxiety and depression.
- Antipsychotics: A specific class of psychiatric drug can be acquired with a prescription for treating psychosis.
- Benzodiazepines: Sedative medications like benzodiazepines slow down both physical and mental functions.
- Cardiac therapy: A program that uses exercises, support, and education for people to get recovered from a heart attack, heart surgery, or other heart problem.
- Antihypertensives: A class of medication used for the treatment of hypertension (high blood pressure).
- Vasodilatadores: Drugs known as vasodilators relax the smooth muscles in the blood vessels, allowing them to expand.
- Beta blockers: The medications used lower blood pressure. The blocks the effects of the hormone epinephrine, more often known as adrenaline.
- Calcium channel blocker: A drug that lowers blood pressure by blocking calcium from entering the heart and artery cells.
- Block SRAA: A system used to reduce blood pressure and slow the progression of renal disease.
- Antilipidemics: Lipid-lowering agents are a type of medicine used in the treatment of high levels of fats.
- Smoking habit: Starting a smoking habit quickly leads to both a psychological and a physiological addiction.
- Hypertension: The medical syndrome known as hypertension causes persistently high blood pressure in the arteries.

III. METHOD

The method section contains the process of data cleaning, the selection of important features, and the proposed model as shown in FIGURE 1. The data cleaning consists of removing the unwanted information, and feature selection consists of a wrapper and filter technique, where the wrapper is accomplished by using SVM along with RFE, and then the MRMR filter is used to avoid feature duplication. The proposed model further consists of the quantum state, data encoding and decoding, and different layers, as briefly discussed below.

A. DATA CLEANING

Data cleaning involves correcting or deleting inaccurate, corrupted, improperly formatted, inoperable, and redundant data from a dataset. When combining different information sources, there are various possibilities for data duplication or labeling errors. In such a context, we have removed 10,463 patients who had no additional diagnostic testing from the dataset, so 5379 (16.1%) out of the 33,372 remaining patients who had CAD during the desired time period are considered. The dataset has been cleaned by removing all the duplicate and irrelevant variables, eliminating nulls and inoperable variables, converting data types, and handling the missing values. Due to null fields, which indicate they do not contain full information for some of the records, the 20 variables that did not match the criterion were deleted.

TABLE 1. Dataset before and after pre-processing.

Parameters	Dataset before cleaning	Dataset after cleaning
Total samples	33372	33372
Features	82	42
Type	Categorical	Categorical
Output	IHD	IHD

TABLE 2. The effective 42 feature after cleaning.

Name of Feature	Treatment	Variable-type and Values
Catheterization	Catheterization procedure	Categorical 1/-1
Ergometry	Ergometry study	Categorical 1/-1
Echo stress	Stress Echocardiography	Categorical 1/-1
Echocardiogram	Echo-cardiogram	Categorical 1/-1
ECG	Electrocardiogram	Categorical 1/-1
Depression	Diagnose depression	Categorical 1/-1
Alcohol	Alcoholism	Categorical 1/-1
Drug dependence	Drug dependence	Categorical 1/-1
Anxiety	Anxiety disorder	Categorical 1/-1
Dementia	Dementia	Categorical 1/-1
Renal insufficiency	Renal failure	Categorical 1/-1
Diabetes mellitus	Insulin	Categorical 1/-1
Diabetes type 1	Type 1 diabetes	Categorical 1/-1
Diabetes type 2	Type 2 diabetes	Categorical 1/-1
Dyslipidemia	Dyslipidemia	Categorical 1/-1
Hypercholesterolemia	Hypercholesterolemia	Categorical 1/-1
Fibrillation-palpitation	Ventricular fibrillation	Categorical 1/-1
Flutter	Auricular flutter	Categorical 1/-1
Heart failure	Cardiac insufficiency	Categorical 1/-1
Obesity	reduced-calorie diet	Categorical 1/-1
Age	Age at notification	Categorical 1/-1
Antithrombotic Agents	antithrombotic agents	Categorical 1/-1
Acetylsalicylic Acid	acetylsalicylic acid	Categorical 1/-1
Proton-pump inhibitors	Proton-pump inhibitors	Categorical 1/-1
Diuretics	Diuretics	Categorical 1/-1
Hormonal contraceptives	Hormonal contraceptives	Categorical 1/-1
Anabolic steroids	Anabolic steroids	Categorical 1/-1
Antigout preparations	Antigout preparations	Categorical 1/-1
Immunomodulators	Immunomodulators	Categorical 1/-1
Antidepressants	Antidepressants	Categorical 1/-1
Antipsychotics	Antypsychotics	Categorical 1/-1
Benzodiazepines	Benzodiazepines	Categorical 1/-1
Cardiac therapy	Cardiac therapy	Categorical 1/-1
Antihypertensives	Antihypertensives	Categorical 1/-1
Vasodilatadores	Vasodilators	Categorical 1/-1
Beta blockers	Beta-blockers	Categorical 1/-1
Calcium-channel blockers	Calcium-channel blockers	Categorical 1/-1
Block SRAA	SRAA blockage	Categorical 1/-1
Antilipidemics	Lipid-lowering	Categorical 1/-1
Smoking habit	Smoking	Categorical 1/-1
Hypertension	Hypertension	Categorical 1/-1

Similarly, in the elimination of duplicate variables, the 14 variables that have 0 values and make no contribution to the model's prediction were also eliminated. Patient admission and discharge dates, patient IDs, and other irrelevant information were also removed since they were not informative. The essential feature between each pair of variables was chosen using a threshold. Subsequently, the clean dataset

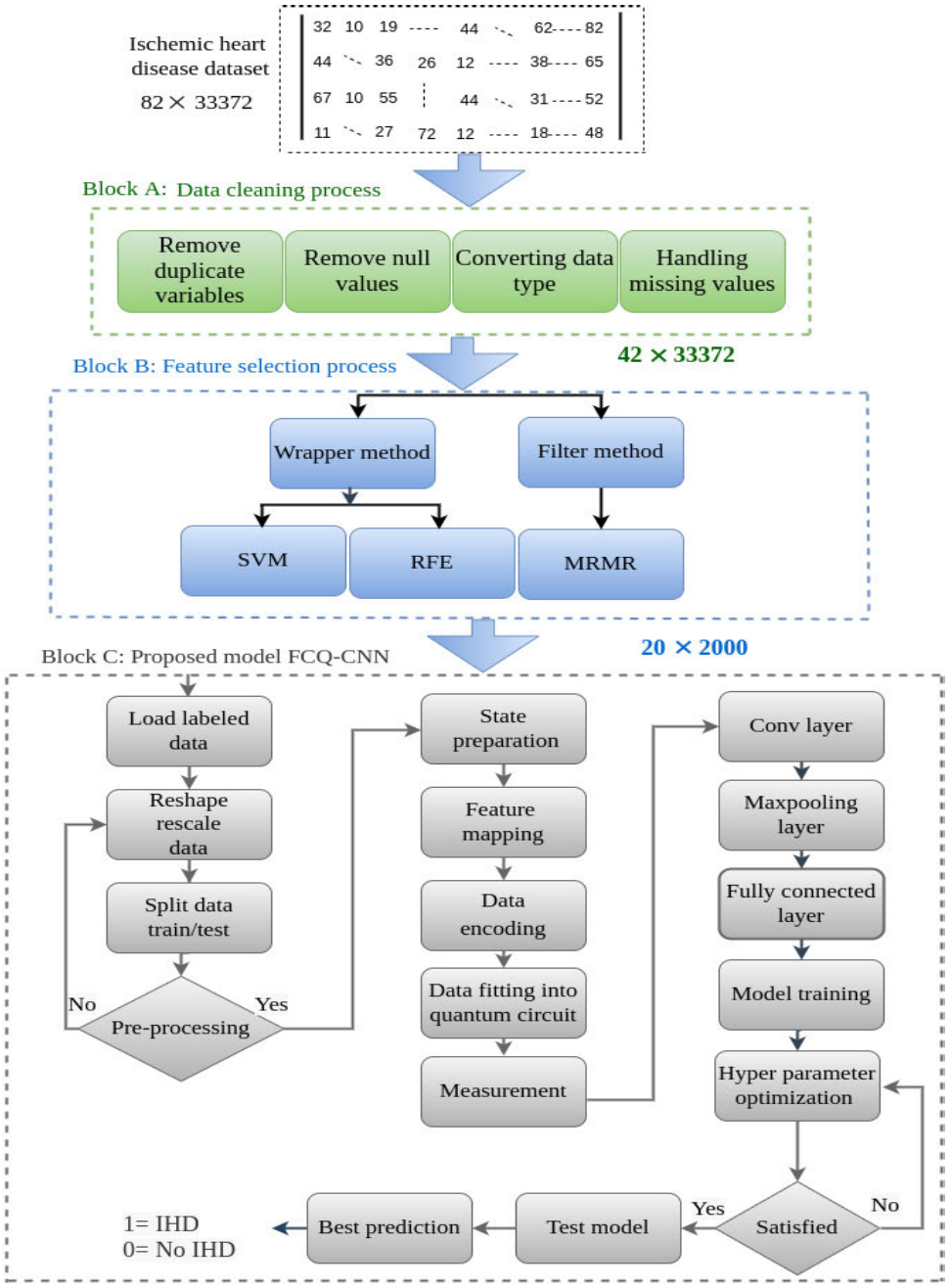


FIGURE 1. The complete block diagram of the proposed methodology.

contains 42 attributes, including the targeted CAD variable as depicted in TABLE 1. The information on the patients who had various therapies, such as cardiac therapy or catheterization, was also included. Moreover, risks associated with various addictions, including alcoholism, certain medicines or pharmaceuticals, including antithrombotic agents, proton pump inhibitors, and anabolic steroids, as well as specific health issues, including diabetes mellitus, obesity, and hypertension, were taken into account. The dataset after the cleaning process, which has 42 effective features, where all the features are categorical, is shown in TABLE 2.

B. FEATURE SELECTION

The technique uses a feature selection procedure to choose the minimum possible input parameters while creating a predictive model. It is preferable to decrease the size of input parameters, which minimizes the computational cost of the model and, under certain conditions, improves its performance. In this research, the SVM and RFE were used in combination with a maximally relevant minimal redundancy (MRMR) filter to minimize the input data dimensions. The goal of the MRMR technique is to choose a group of features that are both highly relevant and minimally redundant for

TABLE 3. Selected 20 important features.

Ranking	Importance	Name of variables
6	0.040610	Catheterization
22	0.013607	Ergometry
18	0.018645	Echo Stress
10	0.024103	Echo-cardiogram
11	0.021145	ECG
13	0.015723	Diabetes millitus
15	0.018850	Hypercholesterolemia
16	0.015983	Fibrillation-palpitation
17	0.014562	Obesity
1	0.285714	Age
3	0.084274	Antithrombotic agents
16	0.015029	Proton pump inhibitors
13	0.016108	Diuretics
31	0.017277	Benzodiazepines
32	0.285714	Cardiac therapy
35	0.089991	Beta blockers
36	0.016288	Calcium channel blockers
37	0.030682	Block SRAA
38	0.076089	Antilipidemics
39	0.016268	Hypertension

identifying different classes of cardiovascular diseases. For an instant, let $D = \{x_{i,k}\}_{n\times K}$ is the expression matrix for the entire dataset, where $x_{i,k}$ is the i_{th} feature of sample k, n represent the total features and K is the total number of samples. So the feature expression for k_{th} sample would be $x_k = (x_{1,k}, x_{2,k}, \ldots, x_{n,k})$, $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,K})$ is i_{th} feature expression, and $G = \{1, 2, \ldots, n\}$ may represent the feature index set. Furthermore, for binary classification, $y_k = \ell \in \{+1, -1\}$ represents the target class values for the k samples by considering $+1$ and -1, respectively. In a subset $S \subset G$, the relevancy R_s of the features is given as [48],

$$R_S = \frac{1}{|S|}\sum_{\ell}\sum_{i\in S} I(\ell, i) \tag{1}$$

$$I(\ell, i) = \sum_{x_i.} p(\ell, x_{i.}) \log \frac{p(\ell, x_i)}{p(\ell)p(x_i)} \tag{2}$$

where $I(\ell, i)$ indicates the mutual information among the class label ℓ and feature i. In a subset S, the mutual information can be used to determine the redundancy of feature i with other features and provide the following value [48].

$$Q_{S,i} = \frac{1}{|S|^2}\sum_{i'\in S, i'\neq i} I\left(i, i'\right) \tag{3}$$

For feature ranking, SVM-RFE was first presented by Guyon et. al [49], which rates the quality performance of the feature subsets for the modeling technique being utilized as a black box evaluator. In essence, the recursive feature elimination technique ranks the features by recursively removing the un-useful features and creating a model based on the remaining ones. The following are the variables of the weight vector w of the SVM that provide the ranking score [50].

$$w = \sum_{k} \alpha_k y_k x_k \tag{4}$$

The MRMR filter may not produce reliable performance when used alone since the classifier operates independently and is not included in the feature selection process. The selected 20 rank wise important features are shown in TABLE 3, where the redundancy between the features is not taken into account. By adding an MRMR filter to SVM-RFE, we hope to improve the model's performance by reducing the number of relevant features that are used more than once. This will also make it easier to classify things.

C. FULLY CONNECTED QUANTUM-CONVOLUTION NEURAL NETWORK (FCQ-CNN)

The Quantum Convolution Neural Network (QCNN) has distinguished itself as one of the machine learning models that perform very well in pattern recognition. The quantum convolution layer can be applied after the data has been encoded into a quantum system. In order to achieve a pure quantum state, a circuit of featuremap has been used for data encoding in the computational basis states, with the number of samples in the probability amplitudes. The subsequent layer, which consists of $O(log(n))$ layers for n input qubits, creates a shallow circuit depth. The desired quantum model structure is capable of avoiding one of the most significant problems connected with parameterized quantum circuit-based algorithms [51]. These architectures seem to be uniquely connected to the tensor network, which offers an easy way to investigate many body physics neural networks, and their interactions. The gradual reduction of the quantum bits is equivalent to the Maxpooling process in CNN. Transnational invariance is a key feature of the QCNN architecture that requires parameterized quantum gate blocks to be similar throughout a layer. The resulting quantum state derived from the I_{th} layer of QCNN can be described as [53].

$$|\psi_i(\boldsymbol{\theta}_i)\rangle\langle\psi_i(\boldsymbol{\theta}_i)| = \mathrm{Tr}_{B_i}\left(U_i(\boldsymbol{\theta}_i)|\psi_{i-1}\rangle\langle\psi_{i-1}|U_i(\boldsymbol{\theta}_i)^{\dagger}\right) \tag{5}$$

The primary objective of the QCNN model is to learn relevant data for accurate prediction of quantum states and their associated labels. At quantum state, the encoded data is denoted by $|x_{in}\rangle$. The features are extracted layer by layer using transformations in parameterized quantum circuits. Quantum measurements on certain qubits are carried out towards the output of the model, which determines the expectation values that represent the classification results as shown in FIGURE 2. The proposed FCQ-CNN model is relatively similar to conventional CNNs, the only difference is that it additionally consists of a quantum layer. The complete diagram of the proposed model is shown in FIGURE 3, which contains quantum layers, convolutional layers, Maxpooling layers, and a fully connected layer.

1) QUANTUM LAYER

The quantum layer contains quantum circuits, quantum gates, and quantum registers, where the quantum parameters are used for computation. The quantum layer further contains a parameterized quantum circuit, quantum feature mapping, and a measurement circuit, as briefly discussed below.

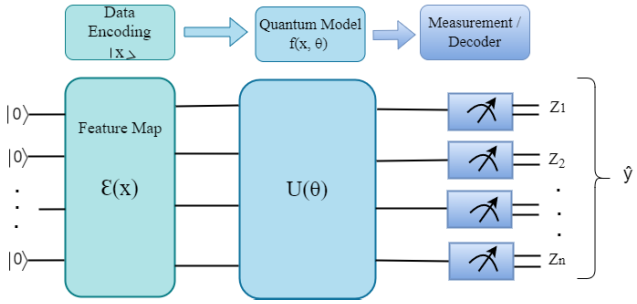


FIGURE 2. Working process of quantum model.

a: PARAMETERIZED QUANTUM CIRCUIT

In a parameterized quantum circuit, real-time classical processing and synchronized quantum operations on quantum data are performed simultaneously. It is composed of several quantum variables, quantum gates, and a measurement circuit that uses the classical data from conventional computers under certain conditions, as shown in FIGURE 4. The qubit, which has two fundamental states denoted by $|1\rangle$ and $|0\rangle$ respectively, is the basic information storage unit. Contrary to conventional bits, which can only accept one value at a time, the qubits, can be in any superposition state [52].

$$|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \in \mathbb{C}^2 \tag{6}$$

In the above equation $\left|\psi^1\right\rangle, \cdots, |\psi^n\rangle$ can be used to represent the quantum states of n qubits in a quantum circuit. The fundamental of the linear space of any n qubit in a quantum state is as, $\{|00\cdots0\rangle, |00\cdots1\rangle, |11\cdots1\rangle\}$. The quantum state $|\Psi\rangle$ in linear space can be described by using the superposition state [52].

$$|\Psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle, \alpha_i \in \mathbb{C} \tag{7}$$

where $|i\rangle$ represents the quantum state defined by the binary form of i. Quantum gates can control the states of qubits in quantum computing. A unitary transformation U, where $UU^\dagger = I$ describes the development of a closed system in quantum theory. A quantum gate that performs a unitary transition U for a quantum system in its initial state $|\psi_0\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle$, acts like matrix-vector multiplication [52].

$$U\,|\psi_0\rangle = U \sum_{i=0}^{2^n-1} \alpha_i |i\rangle = \sum_{i=0}^{2^n-1} \beta_i |i\rangle \tag{8}$$

b: QUANTUM FEATURE MAPPING

A quantum feature map uses a quantum circuit derived from the traditional machine learning kernel approach to encode classical data in the quantum state space. To find a separate hyperplane that classifies nonlinear data, the data are mapped onto a higher-dimensional Hilbert space. Feature map that encodes classical input x_i into a quantum variable $|\Psi(\mathbf{x}_i)\rangle$ by performing ground state $|0\rangle^n$ transformation using n number of unitary gates [53], [54].

$$u_\phi(\mathbf{x}) = U_{\phi(\mathbf{x})} H^{\otimes n} U_{\phi(\mathbf{x})} H^{\otimes n} \tag{9}$$

$$U_{\phi(\mathbf{x})} = \exp\left(i \sum_{S \subseteq [n]} \phi s(\mathbf{x}) \prod_{i \in S} P_i \right) \tag{10}$$

where the Hadamard gate is denoted by H, the subscript $U_{\phi(\mathbf{x})}$ indicates a diagonal gate (unitary gate) on the Paulifeature basis, and (P_i) is the feature space. Finding a separate hyperplane in the new space is necessary since the unitary operations in the initial state pushed the data up to the high dimension of the feature space. The size of the data determines how many qubits are necessary, and unitary gates $U_{\phi(\mathbf{x})}$ are used to encode the data by changing the angle to particular values.

c: QUANTUM MEASUREMENT

The measurement circuit does a critical evaluation to assess the number of classes. The information about the quantum system is not immediately accessible, so we must carry out a quantum measurement to obtain it. This procedure is analogous to taking multiple samples from the distribution of feasible cognitive base states, determining the approximate value, and then assessing the class possibilities by taking a final measurement. For instance, executing a projective measurement on a qubit with Z observable in the state $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$ leads to the generation of -1 and 1, with probabilities of $p(1) = |\alpha|^2$ and $p(-1) = |\beta|^2$. The outcome of the measurement circuit is stochastic, and with each measurement, the quantum state is converted to one of two intermediate states $|0\rangle$ and $|1\rangle$ accordingly. Repeated measurements have been taken to obtain the most precise information possible about the desired outcome. The following is an expression for the expected value of a particular measurement observable output Z of state $|\phi\rangle$ [55].

$$\langle Z \rangle_{|\phi\rangle} \equiv \langle \phi | Z | \phi \rangle = |\alpha|^2 - |\beta|^2 \tag{11}$$

where $Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, such that $\langle Z \rangle \in [-1, 1]$ and $\langle \phi | = (|\phi\rangle)^\dagger$.

2) CONVOLUTIONAL LAYER

During the learning process, a convolutional layer is usually known as the most important component of a CNN. It consists of multiple filters (or kernels), and the parameters must be learned throughout the training process. The filter size is usually smaller than the actual dimension of the input data, and the activation map is created by the convolution of filters with the input volume. The convalutional layer output volume is achieved by overlaying the depth dimension along with each filter activation map. On the other side, the layer defined by the quantum convolution layer has properties similar to those of the classical convolution layer. The quantum convolution layer works with filters in the same way that the traditional

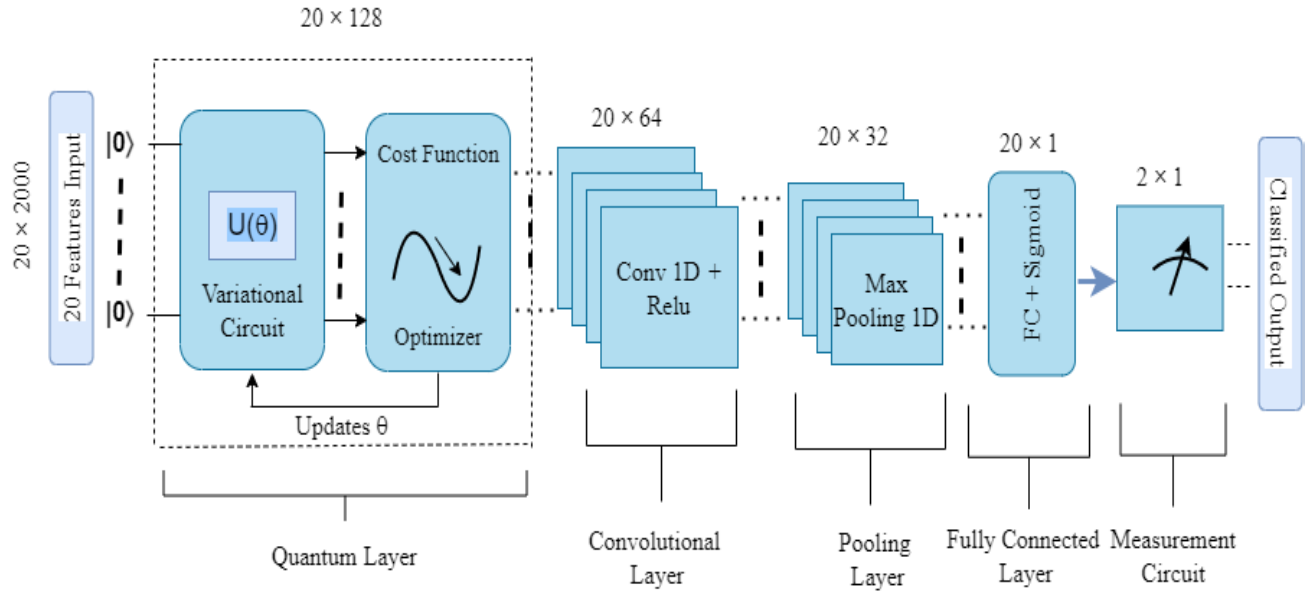


FIGURE 3. The complete FCQ-CNN model's layer-wise overview contains a quantum layer, a convolutional layer, a maxPooling layer, a fully connected layer, and a measurement circuit. The quantum layer further contains a parameterized quantum circuit, quantum feature mapping, and a cost function updated by an optimizer.

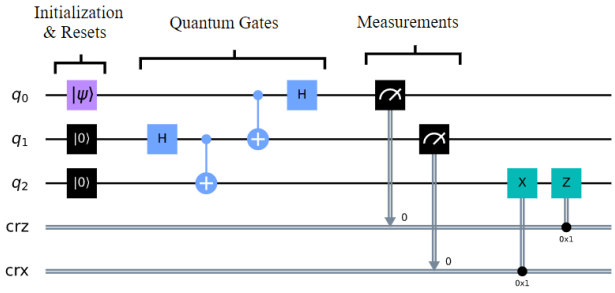


FIGURE 4. Parameterised quantum circuit.

convolution layer does. It applies a filter to the data that comes in and makes feature maps out of new data.

For an instance, consider n number of layers l, if X^l is the input, K^l is the kernel size, and $f : \mathbb{R} \mapsto [0, C]$ with $C > 0$ is a nonlinear function, then the output for l would be $f\left(X^{\ell+1}\right) = f\left(X^{\ell} * K^{\ell}\right)$. The input layer X^l and kernel size K^l stored in the QRAM, and the quantum algorithm is used to obtained a quantum state $\left|f\left(\bar{X}^{\ell+1}\right)\right\rangle$, for $\left\|f\left(\bar{X}^{\ell+1}\right) - f\left(X^{\ell+1}\right)\right\|_{\infty} \leq 2M\epsilon$ for each precision parameter $\epsilon > 0$ and $\eta > 0$, and can be defined as follows:

$$\begin{cases} \left|\mathcal{X}_j^{\ell+1} - f\left(X_j^{\ell+1}\right)\right| \leq 2\epsilon & \text{if} \quad f\left(\bar{X}_j^{\ell+1}\right) \geq \eta \\ \mathcal{X}_j^{\ell+1} = 0 & \text{if} \quad f\left(\bar{X}_j^{\ell+1}\right) < \eta \end{cases} \tag{12}$$

The execution time of the algorithm for each iteration may be,

$$\widetilde{O}\left(\frac{1}{\epsilon\eta^2} \cdot \frac{M\sqrt{C}}{\sqrt{\mathbb{E}\left(f\left(\bar{X}^{\ell+1}\right)\right)}}\right) \tag{13}$$

The above $\mathbb{E}\left(f\left(\bar{X}^{\ell+1}\right)\right)$ is the average value of the quantum state, and $\widetilde{O}$ is the polylogarithmic hide factor that depends on input and kernel size. Furthermore, the kernel size is determined by the Polly logarithmic factor, which allows the QCNN to work with deeper kernels. In forward propagation, the loss function $\mathcal{L}$ is determined by the output of QCNN. For binary classification, the Binary Cross-Entropy has been considered a loss function and may be calculated based on the formula shown below.

$$\mathcal{L} = -\frac{1}{m}\sum_{i}^{m}\left(y_i * \log\left(p\left(y_i\right)\right) + \left(1 - y_i\right) * \log\left(1 - p\left(y_i\right)\right)\right) \tag{14}$$

Here, m shows the number of samples starting from i, the probability of true samples, or class 1 is represented by $p\left(y_i\right)$ and $p\left(1 - y_i\right)$ indicates the probability for class 0 respectively. In the case of back-propagation, a quantum circuit can be successfully characterized with a fully connected quantum neural network that has symbolic similarity with the classical neural network, except that the activation function is not assigned to all nodes. Furthermore, the quantum circuit's input may be one of the 2^n (where n is the number of qubits) having an amplitude of 1. Like traditional deep learning, minimizing error back-propagation is also important for quantum circuits. The gradient of the weights is determined via the back-propagation method by using parameter shift and the chain rule [56], [57] of a partial differential to propagate the gradient back from the network output. Back propagation updates the weights in a manner that reduces the loss by providing the nodes with lower error rates and higher weights, and vice versa. This allows it to learn how much of the overall loss each node is accountable for contributing. The

quantum gates and quantum state $|\psi\rangle$ are represented by complex values for error back-propagation in the simulation of quantum computing. The output $|\psi_{out}\rangle$ may be stated as follows when $|\psi_{in}\rangle$ is the input for n qubits and $W(\theta)$ is the applied quantum circuit parameter network [53].

$$W(\theta)\,|\psi_{\text{in}}\,\rangle = \sum_{j=0}^{2^n-1} c_\theta^j |j\rangle = |\psi_{\text{out}}\,\rangle \tag{15}$$

In above, the state $|j\rangle$ amplitude probability is denoted by c_θ^j and $\left|c_\theta^j\right|^2 = p_\theta^j$ shows the observation probability of state $|j\rangle$.

3) POOLING LAYER

Pooling layers seeks to continuously reduce the dimensionality of the representation, which reduces the computational cost and number of parameters for the model. In input, the pooling layer scales the dimension of every activation map by using the "MAX" function. The majority of CNNs apply 2-by-2-dimensional kernels in steps of 2 along the spatial dimensions of the input using max-pooling layers. This keeps the depth volume at its actual size while scaling the activation map down to 25% of its original size. Due to the pooling layer's destructive characteristics, there are two frequently utilized methods for max-pooling. The stride and filters of the pooling layer are usually set to 3 and 3, respectively, allowing the layer to extend throughout the maximum range of input spatial dimensions. Additionally, overlapping pooling may be used with a stride and kernel size of 2 and 3, respectively. Because pooling is destructive, if the kernel size is greater than 3, the model's performance will usually go down a lot.

4) FULLY CONNECTED LAYER

The output from the last pooling or convolutional layer is passed through to the fully connected layer after being flattened. The fully connected layer is composed of a collection of dependent nonlinear functions, each of which is made up of a perceptron or neuron. The neurons perform a linear transformation on the input vector using a weight matrix. After the several convolutional and pooling layers in a quantum environment, the amount of qubits has decreased with the implementation of the FC layer. The desired qubit circuit has been used in the case of a relatively small system to accomplish the classification. The highly entangling circuits, which include CNOT gates and universal single-qubit gates with basic cells composed of $n(n+1)$ and $n(n-1)$, are considered for the FC layer. Each CNOT has a variable indicating whether it is active or inactive and working as a unitary gate [58]. The two-qubit universal gate operation is shown in FIGURE 5.

To obtain the prediction values, the FCQ-CNN model last step involves measuring a certain number of output qubits. The measurement circuit converts the quantum variables $|0\rangle$ and $|1\rangle$ back into binary variables, 0 and 1. The mapping of expected values to classification outputs may be carried out in various ways. Formally, the FCQ-CNN model's output for the input *xin* is expressed as $f\,(\boldsymbol{\theta}, \boldsymbol{x}_{in})$. In general, measuring one qubit and using the expected value as the output $f\,(\boldsymbol{\theta}, \boldsymbol{x}_{in}) \equiv \langle Z\rangle$ is a easy way for binary classification tasks. Then, as a result, $\langle Z\rangle \geq 0$ represents the classification of samples into class 1, and $\langle Z\rangle < 0$ denotes classifying the sample into class 0.

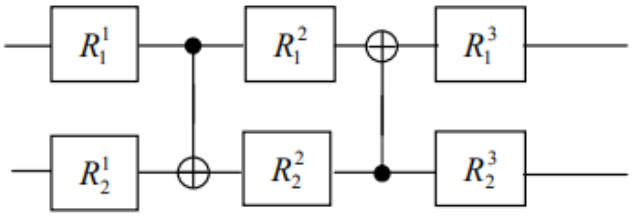


FIGURE 5. Two qubits universal gate operation.

IV. RESULTS

This section contains the experimental results that demonstrate how the actual quantum convolutional neural network performs when applied to classify diseases using the cardiovascular dataset. The TensorFlow Quantum platform is used for the simulation of the proposed model. Initially, there are 42 attributes in the cardiovascular dataset that are associated with IHD. However, due to resource constraints faced by the quantum computing simulation environment, the features are limited to 20 features, and the total number of samples considered for the desired experiment is 2000. The Torch interface, Qiskit, Keras, and Cirq libraries are used to implement the models in the Python 3 environment. An application programming interface (API) is used to run the proposed FCQ-CNN model on the IBM state-vector simulator. The features are encoded by using a PauliFeature map, where the layer of the Hadamard gate along with entangling blocks maps the classical data into quantum variables. The quantum circuit is rotated with a depth of 10, and a measurement circuit is used to decode the data again into classical 0,1. All models use a Relu activation function for the input and hidden layers, while a Sigmoid function is used for binary classification on the output layer.

A. HYPERPARAMETER TUNING

It is possible for neural networks to learn complex correlations among their inputs and outputs. Although many of these correlations will be visible during training but not in actual test data because of sampling noise. This problem could result in overfitting, which would decrease the model's capabilities for prediction. In short, it is quite difficult to avoid the overfitting issue and obtain the generalized prediction of the proposed model. In this paper, we applied several optimization techniques for hyperparameters to get the best possible set. The input dataset is first normalized by using min-max and standard scalars, which scale the data in a defined range of 1 and 0. The data are then encoded using different feature maps, such as the PauliFeature map, ZFeature map, and ZZFeature map, with circuit depths

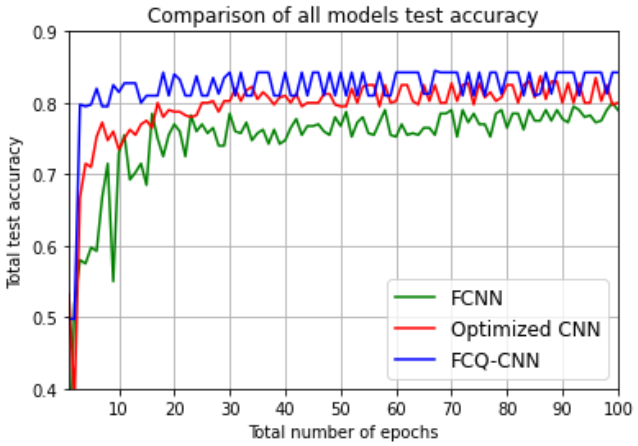


FIGURE 6. Comparison of testing accuracy of all models.

TABLE 4. Set of optimal hyperparameters of each model.

Optimal hyperparameters	FCNN model	Optimized-CNN model	Proposed model
Learning rate	0.001	0.001	0.001
Optimizer	Adam	Adam	Adam
Batch size	32	32	32
Epochs	100	60	20
Loss function	Binary crossentropy	Binary crossentropy	Binary crossentropy
Activation-function for dense layer	Relu	Relu	Relu
Activation-function for output layer	Sigmoid	Sigmoid	Sigmoid

of 5, 10, and 15, where the PauliFeature map with a circuit depth of 10 was found to be the best for our model. The model has been evaluated by using various loss functions, and for our binary classification task, Binary Cross-Entropy produced promising results. The effective epochs for the FCNN, Optimized-CNN, and FCQ-CNN models are 100, 60, and 20, respectively. This is because beyond a certain point, the models' accuracy does not increase and remains constant, as shown in FIGURE 6. During the training process, Adam (adaptive optimization techniques) is used for 100 epochs, which outperformed stochastic gradient descent (SGD) in terms of optimal performance. The model is initially run using three different learning rates (0.01, 0.001, and 0.0001) with separate batch sizes (16, 32, and 64), and it gets good results by considering learning rate 0.001 and batch size 32. The number of filters for each layer is 128, 64, and 32, respectively, where each filter size is 3. The input and hidden layers are equipped with the Relu activation function, and the output layer uses the Sigmoid function. The model is trained with the training dataset, the weights with the lowest training loss are saved, and the prediction performance on the labeled test dataset is assessed. Finally, the desired optimal weights are used for class prediction of the test dataset. TABLE 4 indicates the optimal hyperparameter values used for each model.

TABLE 5. Comparisons of FCQ-CNN with Optimized-CNN and FCNN models.

Parameters	FCNN model	Optimized-CNN model	Proposed model
Dataset	Cardiovascular	Cardiovascular	Cardiovascular
Total samples	2000	2000	2000
Total test samples	400	400	400
Total train samples	1600	1600	1600
No of layers	4	5	5
Testing accuracy	79.2%	80.4%	**84.6%**
Testing loss	0.4	0.3	**0.28**

B. COMPARISON OF FCQ-CNN MODEL WITH FCNN AND OPTIMIZED-CNN MODELS

The competency of the proposed FCQ-CNN model is obtained by comparing it with the classically optimized-CNN and FCNN models by considering the cardiovascular dataset in terms of testing accuracy and testing loss, respectively, as shown in TABLE 5. The classification accuracy of each model is obtained by considering the same hyperparameter optimization, which also ensures a fair comparison. It is important to keep a small number of nodes in the input layer of all models to be trained properly with a limited number of parameters. The FCNN model contains one input layer with 20 nodes, two dense layers with 32 and 16 nodes, and an output layer with 1 node for binary classification, respectively. Likewise, the optimized-CNN contains the input layer, two convolutional layers, and two maxpooling layers with the same number of nodes. In the case of FCQ-CNN, after the quantum layer there is a convolutional layer, a maxpooling layer, and a fully connected layer. The input layer, the hidden layer, and the convolutional layer of all models use the Relu activation functions, while the output layer uses the sigmoid function, which converts any input to an output between 0 and 1 for the classification of cardiovascular diseases. The output of the model sigmoid predicts one value for class 0, if $\hat{y} < 0.5$ and for class 1, if $\hat{y} > 0.5$, where $\hat{y}$ is the predicted output of the model. The model is trained with 1600 (80%) samples out of a total of 2000 samples, and 400 (20%) samples are used for testing purposes. An Adam optimizer is used, where the learning rate is considered to be 0.001 for each model. FIGURE 6 shows the testing accuracy of FCNN, Optimized-CNN, and FCQ-CNN models, where the highest testing accuracy achieved by FCQ-CNN is 84.6%, Optimized-CNN accuracy is 80.4% and FCNN has 79.2% accuracy, respectively. Likewise, the calculated loss comparison is indicated in FIGURE 7, which also shows that the proposed FCQ-CNN model has the lowest training loss and has a value of 0.24, the optimized-CNN has a loss value of 0.3 and the FCNN model has a calculated loss of 0.4 respectively. In the comparison of the Optimized-CNN and FCNN models, the reason for the highest accuracy and low error rate of Optimized-CNN is due to the selection of the best subset of hyperparameters. Moreover, the Optimized-CNN model seems more robust than the FCNN model because of its shared-weights architecture, comparatively small set of parameters, and characteristics

TABLE 6. Comparisons of FCQ-CNN with previous published quantum models.

Parameters	Maheshwari *et.al*	PennyLane team *et.al*	Yano *at.al*	Kerenidis *at.al*	Proposed model
Dataset	Synthetic, Sonar, Diabetes	Iris	Breast-Cancer, Heart Disease	MNIST	Cardiovascular Disease
Used Model	QSVM, VQC	QRF	VQC, VQC+QRAC	QCNN	FCQ-CNN
Total features	10, 10,13	10	9, 13	8	20
Number of classes	2,2,2	3	2, 2	2	2
Training samples	400, 166, 2121	125	146, 243	60,000	1600
Testing samples	100, 42, 531	25	50, 60	10,000	400
Total size	500, 208, 2651	150	196, 303	70,000	2000
Model accuracy	94%, 75%, 76%, 71%,-74.5%, 69%	72%	69%-73%, 82%, 85%	82.8%	**84.6%**

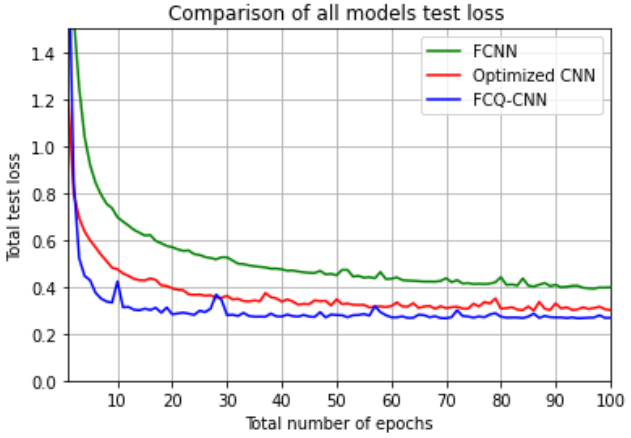


FIGURE 7. Comparison of testing loss of all models.

of translation invariance for binary classification problems. In the FCNN model, each node in the dense and flattened layers is connected in the network by utilizing the most trainable parameters, which reduces the model's performance and increases its computational cost. On the other hand, the FCQ-CNN has the highest test accuracy and lowest test loss of the two models (optimized-CNN and FCNN). The FCQ-CNN model loss seems more stable after 30 epochs, this is due to the quantum network's quick learning convergence, which requires a comparatively smaller iteration to achieve the flattening of the loss function and accuracy. The testing and training times cannot be compared, because simulating a quantum network naturally takes a long time and demands the numerical equivalent of quantum computation. In the end, the quantum algorithm does a better job at IHD classification tasks than a classical algorithm and offering substantial advantages for a large amount of data.

V. DISCUSSION

The proposed FCQ-CNN is also compared with recently published quantum models, such as the Quantum Support Vector Machine (QSVM), the Quantum Random Forest (QRF), the Quantum K-Nearest Neighbor (QKNN), and the Quantum Convolutional Neural Network (QCNN). The TABLE 6 shows that the proposed FCQ-CNN is more competent than the existing quantum models in terms of high test accuracy, least complexity, and low computational loss. Numerous studies on quantum-enhanced methods have been carried out to tackle issues with machine learning. For an instant [59], presented various QML approaches, i.e., QSVM, VQC, and amplitude-encoding VQC, by considering the three different datasets. The highest accuracy for the UCI diabetes dataset is achieved at 74.1%, 68.7%, and 74.5% for QSVM, VQC, and amplitude-encoding VQC, respectively. The algorithms are executed by using the IBM simulator and the Qiskit framework [34]. Similarly, comparing our results to those of a literature review of the ensemble learning approach proposed by the Pennylane project team [60]. It uses two different Quantum Processing Units (QPUs), forest.qvm and qiskit.aer, each working as a separate cluster of trees. For identifying three different groups, 150 samples from the iris dataset were employed. A PauliFeature map is used to convert the classical data, and RX gate rotation is used to determine the depth of the circuit. For both the training and testing datasets, the ensemble model predicted results are recorded as 83.2 percent and 72%, respectively. The model's reliability is verified by comparing it with Yano et.al [61], which exploits quantum random access coding to effectively map such discrete features into a finite amount of qubits for VQC. The goal of the work is to show that QRAC can speed up VQC training by lowering its settings and minimizing the number of qubits required for the mapping. The models are evaluated using UCI breast cancer and heart disease data with 286 instances and 9 features and 303 instances and 13 features, respectively. The models work well and obtained the maximum accuracy of 72% and 88% for VQC and VQC-QRAC for both datasets. The quantum CNN completely replicates the classical CNN by supporting non-linearities and pooling operations. A new quantum approach with ℓ_∞ norm and new quantitative research algorithms in the context of information processing is introduced by Kerenidis et.al [62]. For classification, the MNIST datasets' classification using numerical simulations has been considered, which demonstrates the effectiveness of the QCNN in a real-world scenario. The model achieved the highest test accuracy 82.8% and

74.8% and error rates of 0.51 and 0.77 with $\ell = 0.01, 0.1$ respectively.

In summary, our proposed FCQ-CNN model offers significant results on the cardiovascular dataset and is competitive with recently published different quantum machine learning models [59], [60], [61], [62].

VI. CONCLUSION

Quantum machine learning and data science have interesting future uses thanks to fully parameterized quantum convolutional neural networks. In this paper, a comprehensive benchmark of the FCQ-CNN is presented for handling the classification tasks on classical data. The FCQ-CNN algorithm can be customized using a variety of factors, including the design of parameterized quantum circuits, quantum filters, pooling operators, classical data preprocessing techniques, quantum feature mapping, optimizer, and cost functions. The data used for this experiment comes from public hospital and is based on healthcare records relating to heart disease. In the first stage, the data is cleaned by eliminating any redundant and irrelevant variables, nullifying inoperative variables, converting data types, and handling missing values. Combining a support vector machine (SVM) with recursive feature elimination (RFE) as a wrapper, and a maximally relevant minimal redundancy (MRMR) filter, the important 20 features are selected. The input classical data has been converted into quantum variables using quantum feature mapping (PauliFeature map). Using the IBM state-vector simulator, the model is tested with 20 qubits of QCNN for the binary classification of cardiovascular datasets, which reduces the number of free parameters. The generalized FCQ-CNN architecture put forward in this study has an input layer with 20 nodes and two quantum filters. For the input layer and convolutional layer, a Relu activation function is used, while the output layer has a sigmoid function to accomplish binary classification. A measurement circuit is used at the output layer to decode the quantum variables. The model has been tested using several optimizers and learning rates, and it produced statistically significant results using the Adam optimizer and a learning rate of 0.001. Despite the limited number of free parameters, the FCQ-CNN model achieved the highest testing accuracy of 84.6% in all instances for cardiovascular datasets. The proposed model results were also compared to the optimized CNN and FCNN models, and it was noticed that FCQ-CNN performed substantially better than both classical models under the same training conditions for the same data samples.

The primary limitation of this work is the use of a small number of data samples compared to a small number of qubits. The quantum computing device requires more logic gates to use larger datasets and more qubits, which increases the computational cost and prolongs the model execution time. These limitations might have an impact on quantum states, where a wrong rotation might cause an error in the result. In the future, a universally fault-tolerant quantum computer will be needed to efficiently handle issues like integer factorization and unstructured database search. This computer will need millions of low-error, long-coherence-time qubits.

ACKNOWLEDGMENT

The Clinical Research Ethics Committee of Euskadi (PI202031) in Spain validated the research protocols. The article reflects only the author's view, and the Agency is not responsible for any use that may be made of the information it contains.

REFERENCES

[1] G. Renugadevi, G. A. Priya, B. D. Sankari, and R. Gowthamani, "Predicting heart disease using hybrid machine learning model," *J. Phys., Conf.*, vol. 1916, no. 1, May 2021, Art. no. 012208, doi: 10.1088/1742-6596/1916/1/012208.

[2] R. Das, I. Turkoglu, and A. Sengur, "Effective diagnosis of heart disease through neural networks ensembles," *Exp. Syst. Appl.*, vol. 36, no. 4, pp. 7675–7680, 2009, doi: 10.1016/j.eswa.2008.09.013.

[3] S. Nyamathulla and R. Varikuti, "Predicting heart disease with hybrid machine learning algorithms," *SSRN Electron. J.*, May 2022, Art. no. 4121828, doi: 10.2139/ssrn.4121828.

[4] D. C. Yadav and S. Pal, "Prediction of heart disease using feature selection and random forest ensemble method," *Int. J. Pharm. Res.*, vol. 12, no. 4, pp. 56–66, 2020, doi: 10.31838/ijpr/2020.12.04.013.

[5] M. S. Oh and M. H. Jeong, "Sex differences in cardiovascular disease risk factors among Korean adults," *Korean J. Med.*, vol. 95, no. 4, pp. 266–275, Aug. 2020, doi: 10.3904/kjm.2020.95.4.266.

[6] P. Jousilahti, E. Vartiainen, J. Tuomilehto, and P. Puska, "Sex, age, cardiovascular risk factors, and coronary heart disease," *Circulation*, vol. 99, no. 9, pp. 1165–1172, Mar. 1999, doi: 10.1161/01.cir.99.9.1165.

[7] (Jun. 11, 2021). *Cardiovascular Diseases (CVDS)*. World Health Organization. Accessed: Oct. 17, 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[8] K. Uyar and A. Ilhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks," *Proc. Comput. Sci.*, vol. 120, pp. 588–593, Jan. 2017, doi: 10.1016/j.procs.2017.11.283.

[9] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, I. Garciá-Magarinõ, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," *Mobile Inf. Syst.*, vol. 2018, pp. 1–21, Dec. 2018, doi: 10.1155/2018/3860146.

[10] V. Biksham, V. Srujana, I. Meghana, B. Harshath, and G. Tarun, "Heart disease prediction using machine learning," *YMER*, vol. 21, no. 4, pp. 489–494, Apr. 2022, doi: 10.37896/YMER21.04/48.

[11] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–11, Jul. 2021, doi: 10.1155/2021/8387680.

[12] A. S.-Y. Kong, K.-S. Lai, S.-H.-E. Lim, S. Sivalingam, J.-Y. Loh, and S. Maran, "MiRNA in ischemic heart disease and its potential as biomarkers: A comprehensive review," *Int. J. Mol. Sci.*, vol. 23, no. 16, p. 9001, Aug. 2022, doi: 10.3390/ijms23169001.

[13] I. Kononenko, "Machine learning for medical diagnosis: History, state of the art and perspective," *Artif. Intell. Med.*, vol. 23, pp. 89–109, Aug. 2001, doi: 10.1016/S0933-3657(01)00077-X.

[14] F. M. J. M. Shamrat, M. A. Raihan, A. K. M. S. Rahman, I. Mahmud, and R. Akter, "An analysis on breast disease prediction using machine learning approaches," *Int. J. Sci. Technol. Res.*, vol. 9, no. 2, pp. 2450–2455, Feb. 2020.

[15] J. Mackay, G. Mensah, and K. Greenland. (Jan. 1, 2004). *The Atlas of Heart Disease and Stroke/Judith Mackay and George Mensah; With Shanthi Mendis and Kurt Greenland*. World Health Organization. Accessed: Oct. 17, 2022. [Online]. Available: https://apps.who.int/iris/handle/10665/43007

[16] M. S. Amin, Y. K. Chiam, and K. D. Varathan, "Identification of significant features and data mining techniques in predicting heart disease," *Telematics Informat.*, vol. 36, pp. 82–93, Mar. 2019, doi: 10.1016/j.tele.2018.11.007.

[17] N. Kausar, S. Palaniappan, B. B. Samir, A. Abdullah, and N. Dey, "Systematic analysis of applied data mining based optimization algorithms in clinical attribute extraction and classification for diagnosis of cardiac patients," *Intell. Syst. Ref. Libr.*, vol. 96, pp. 217–231, Jan. 2016, doi: 10.1007/978-3-319-21212-8_9.

[18] D. Jain and V. Singh, "Feature selection and classification systems for chronic disease prediction: A review," *Egyptian Informat. J.*, vol. 19, no. 3, pp. 179–189, Nov. 2018, doi: 10.1016/j.eij.2018.03.002.

[19] J. Neumann, C. Schnörr, and G. Steidl, "Combined SVM-based feature selection and classification," *Mach. Learn.*, vol. 61, nos. 1–3, pp. 129–150, 2005, doi: 10.1007/s10994-005-1505-9.

[20] M. Ashraf, S. M. Ahmad, N. A. Ganai, R. A. Shah, M. Zaman, S. A. Khan, and A. A. Shah, "Prediction of cardiovascular disease through cutting-edge deep learning technologies: An empirical study based on TENSORFLOW, PYTORCH and KERAS," *Adv. Intell. Syst. Comput.*, vol. 1165, pp. 239–255, Aug. 2021, doi: 10.1007/978-981-15-5113-0_18.

[21] X. Liu, X. Wang, Q. Su, M. Zhang, Y. Zhu, Q. Wang, and Q. Wang, "A hybrid classification system for heart disease diagnosis based on the RFRS method," *Comput. Math. Methods Med.*, vol. 2017, pp. 1–11, Jan. 2017, doi: 10.1155/2017/8272091.

[22] N. Mathur, J. Landman, Y. Yvonna Li, M. Strahm, S. Kazdaghli, A. Prakash, and I. Kerenidis, "Medical image classification via quantum neural networks," 2021, *arXiv:2109.01831*.

[23] T. A. Shaikh and R. Ali, "Quantum computing in big data analytics: A survey," in *Proc. 16th IEEE Int. Conf. Comput. Inf. Technol. CIT, 6th Int. Symp. Cloud Serv. Comput. IEEE SC2 Int. Symp. Secur. Priv. Soc. Networks Big Data, Soc.* no. 4, pp. 112–115, Dec. 2017, doi: 10.1109/CIT.2016.79.

[24] H. Gupta, H. Varshney, T. K. Sharma, N. Pachauri, and O. P. Verma, "Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction," *Complex Intell. Syst.*, vol. 8, no. 4, pp. 3073–3087, Aug. 2022, doi: 10.1007/s40747-021-00398-7.

[25] Z. Hameed, B. Garcia-Zapirain, J. J. Aguirre, and M. A. Isaza-Ruget, "Multiclass classification of breast cancer histopathology images using multilevel features of deep convolutional neural network," *Sci. Rep.*, vol. 12, no. 1, pp. 1–21, Sep. 2022, doi: 10.1038/s41598-022-19278-2.

[26] D. Sierra-Sosa, J. Arcila-Moreno, B. Garcia-Zapirain, C. Castillo-Olea, and A. Elmaghraby, "Dementia prediction applying variational quantum classifier," 2020, *arXiv:2007.08653*.

[27] D. Maheshwari, B. Garcia-Zapirain, and D. Sierra-Sosa, "Quantum machine learning applications in the biomedical domain: A systematic review," *IEEE Access*, vol. 10, pp. 80463–80484, 2022, doi: 10.1109/ACCESS.2022.3195044.

[28] J. Amin, M. Sharif, N. Gul, S. Kadry, and C. Chakraborty, "Quantum machine learning architecture for COVID-19 classification based on synthetic data generation using conditional adversarial neural network," *Cognit. Comput.*, vol. 14, no. 5, pp. 1677–1688, Sep. 2022, doi: 10.1007/s12559-021-09926-6.

[29] B. K. Mishra and I. Global, *Technology Road Mapping for Quantum Computing and Engineering*. Hershey, PA, USA: IGI Global, 2022.

[30] A. Perdomo-Ortiz, M. Benedetti, J. Realpe-Gómez, and R. Biswas, "Opportunities and challenges for quantum-assisted machine learning in near-term quantum computers," *Quantum Sci. Technol.*, vol. 3, no. 3, pp. 1–13, 2018, doi: 10.1088/2058-9565/aab859.

[31] G. Li, Z. Song, and X. Wang, "VSQL: Variational shadow quantum learning for classification," 2020, *arXiv:2012.08288*.

[32] J. Qi, C.-H. H. Yang, and P.-Y. Chen, "QTN-VQC: An end-to-end learning framework for quantum neural networks," 2021, *arXiv:2110.03861*.

[33] I. Rojas, G. Joya, and A. Catala, "Advances in computational intelligence," in *Proc. 13th Int. Work-Conf. Artif. Neural Netw. (IWANN)*, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 9095. Palma De Mallorca: Spain, 2015, pp. 276–289, doi: 10.1007/978-3-319-19222-2.

[34] S. A. Alasadi and W. S. Bhaya, "Review of data preprocessing techniques in data mining," *J. Eng. Appl. Sci.*, vol. 12, no. 16, pp. 4102–4107, 2017.

[35] P. Mishra, A. Biancolillo, J. M. Roger, F. Marini, and D. N. Rutledge, "New data preprocessing trends based on ensemble of multiple preprocessing techniques," *TrAC Trends Anal. Chem.*, vol. 132, Nov. 2020, Art. no. 116045, doi: 10.1016/j.trac.2020.116045.

[36] S. A. Billings and K. L. Lee, "Nonlinear Fisher discriminant analysis using a minimum squared error cost function and the orthogonal least squares algorithm," *Neural Netw.*, vol. 15, no. 2, pp. 263–270, 2002, doi: 10.1016/S0893-6080(01)00142-3.

[37] A. Khabbazi, E. Atashpaz-Gargari, and C. Lucas, "Imperialist competitive algorithm for minimum bit error rate beamforming," *Comput. Eng.*, vol. 1, pp. 125–133, Jan. 2009.

[38] D. D. Miller and E. W. Brown, "Artificial intelligence in medical practice: The question to the answer?" *Amer. J. Med.*, vol. 131, no. 2, pp. 129–133, Feb. 2018, doi: 10.1016/j.amjmed.2017.10.035.

[39] N. Mishra, A. Bisarya, S. Kumar, B. K. Behera, S. Mukhopadhyay, and P. K. Panigrahi, "Cancer detection using quantum neural networks: A demonstration on a quantum computer," 2019, *arXiv:1911.00504*.

[40] Y. Bengio and Y. Lecun, "Scaling learning algorithms toward AI," *Large-Scale Kernel Mach.*, vol. 35, no. 1, pp. 1–41, Aug. 2019, doi: 10.7551/mitpress/7496.003.0016.

[41] H. Yang, L. Lu, and W. Zhou, "A novel optimization sizing model for hybrid solar-wind power generation system," *Sol. Energy*, vol. 81, no. 1, pp. 76–84, 2007, doi: 10.1016/j.solener.2006.06.010.

[42] T. M. Khan and A. Robles-Kelly, "A derivative-free method for quantum perceptron training in multi-layered neural networks," in *Neural Information Processing* (Communications in Computer and Information Science), vol. 1333. Cham, Switzerland: Springer, 2020, pp. 241–250, doi: 10.1007/978-3-030-63823-8_29.

[43] F. Arute et al., "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019, doi: 10.1038/s41586-019-1666-5.

[44] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, "Characterizing quantum supremacy in near-term devices," *Nature Phys.*, vol. 14, pp. 595–600, Apr. 2018, doi: 10.1038/s41567-018-0124-x.

[45] V. Dunjko and H. J. Briegel, "Machine learning & artificial intelligence in the quantum domain: A review of recent progress," *Rep. Prog. Phys.*, vol. 81, no. 7, Jul. 2018, Art. no. 074001, doi: 10.1088/1361-6633/aab406.

[46] N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd, "Continuous-variable quantum neural networks," *Phys. Rev. Res.*, vol. 1, no. 3, pp. 1–22, Oct. 2019, doi: 10.1103/PhysRevResearch.1.033063.

[47] A. G. Olea, M. J. Acosta, I. D. Gonzalez, M. B. G. Zapirain, I. F. De La Prieta, M. M. Rada, J. M. O. Merodio, U. I. Rodriguez, I. P. Lili, A. R. Rodriguez, and E. A. De Luis, "Electronic health records features associated to coronary artery disease in patients with chest pain," *Eur. Heart J.*, vol. 42, no. 1, Oct. 2021, Art. no. ehab724.1151, doi: 10.1093/eurheartj/ehab724.1151.

[48] P. A. Mundra and J. C. Rajapakse, "SVM-RFE with MRMR filter for gene selection," *IEEE Trans. Nanobiosci.*, vol. 9, no. 1, pp. 31–37, Mar. 2009, doi: 10.1109/TNB.2009.2035284.

[49] H. A. Le Thi, V. V. Nguyen, and S. Ouchani, "Gene selection for cancer classification using DCA," in *Advanced Data Mining and Applications* (Lecture Notes in Computer Science), vol. 5139. Berlin, Germany: Springer, 2008, pp. 62–72, doi: 10.1007/978-3-540-88192-6_8.

[50] H. Sanz, C. Valim, E. Vegas, J. M. Oller, and F. Reverter, "SVM-RFE: Selection and visualization of the most relevant features through non-linear kernels," *BMC Bioinf.*, vol. 19, no. 1, pp. 1–18, Dec. 2018, doi: 10.1186/s12859-018-2451-4.

[51] M. Watabe, K. Shiba, C.-C. Chen, M. Sogabe, K. Sakamoto, and T. Sogabe, "Quantum circuit learning with error backpropagation algorithm and experimental implementation," *Quantum Rep.*, vol. 3, no. 2, pp. 333–349, May 2021, doi: 10.3390/quantum3020021.

[52] G. Chen, Q. Chen, S. Long, W. Zhu, Z. Yuan, and Y. Wu, "Quantum convolutional neural network for image classification," *Pattern Anal. Appl.*, pp. 1–13, Sep. 2022, doi: 10.1007/s10044-022-01113-z.

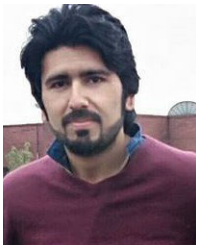
[53] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, pp. 209–212, Mar. 2019, doi: 10.1038/s41586-019-0980-2.

[54] D. Sierra-Sosa, J. D. Arcila-Moreno, B. Garcia-Zapirain, and A. Elmaghraby, "Diabetes Type 2: Poincaré data preprocessing for quantum machine learning," *Comput. Mater. Contin.*, vol. 67, no. 2, pp. 1849–1861, 2021, doi: 10.32604/cmc.2021.013196.

[55] Y. Wang, Y. Wang, C. Chen, R. Jiang, and W. Huang, "Development of variational quantum deep neural networks for image recognition," *Neurocomputing*, vol. 501, pp. 566–582, Aug. 2022, doi: 10.1016/j.neucom.2022.06.010.

[56] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware," *Phys. Rev. A, Gen. Phys.*, vol. 99, no. 3, pp. 1–8, Mar. 2019, doi: 10.1103/PhysRevA.99.032331.

[57] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Phys. Rev. A, Gen. Phys.*, vol. 98, no. 3, pp. 1–3, Sep. 2018, doi: 10.1103/PhysRevA.98.032309.
[58] P. B. De Sousa and R. V. Ramos, "The general decomposition of single-qubit, two-qubits and three-qubits quantum gates," Cornell Univ., Ithaca, NY, USA, Tech. Rep., 2006.
[59] D. Maheshwari, D. Sierra-Sosa, and B. Garcia-Zapirain, "Variational quantum classifier for binary classification: Real vs synthetic dataset," *IEEE Access*, vol. 10, pp. 3705–3715, 2022, doi: 10.1109/ACCESS.2021.3139323.
[60] PennyLane Dev Team. (2021). *Ensemble Classification With Forest and Qiskit Devices¶, Ensemble Classification With Forest and Qiskit Devices-PennyLane Documentation*. Accessed: Oct. 17, 2022. [Online]. Available: https://pennylane.ai/qml/demos/ensemble-multi-qpu.html
[61] H. Yano, Y. Suzuki, R. Raymond, and N. Yamamoto, "Efficient discrete feature encoding for variational quantum classifier," in *Proc. IEEE Int. Conf. Quantum Comput. Eng. (QCE)*, Oct. 2020, pp. 11–21, doi: 10.1109/QCE49297.2020.00012.
[62] I. Kerenidis, J. Landman, and A. Prakash, "Quantum algorithms for deep convolutional neural networks," 2019, *arXiv:1911.01117*.



UBAID ULLAH received the B.Sc. and M.Sc. degrees in electronics from the Department of Electronics, University of Peshawar, Pakistan, in 2013 and 2016, respectively, and the M.Phil. degree in electronics from the Department of Electronics, Quaid-i-Azam University, Islamabad, Pakistan, in 2018. Currently, he is working as a Research Assistant under the COFUND Marie Skłodowska-Curie Fellowship at the Faculty of Engineering, University of Deusto, Bilbao, Spain. He has published many research articles in various journals, such as IEEE Access, *Energies* (MDPI), and *Symmetry* (MDPI). His current research interest includes quantum machine learning in healthcare domain. He has also served as a reviewer for various journals.

ALAIN GARCíA OLEA JURADO is currently a Cardiologist at Basurto University Hospital and belongs to the Biocruces Clinical Cardiology Research Unit. He has also carried out review tasks for Elsevier, being among the reviewers of the book titled *Netter's Introduction to Clinical Procedures* (1st Edition). In the field of artificial intelligence research, he has presented papers at national and European congresses on the prediction of ischemic heart disease and estimation of atrial fibrillation recurrence from the electronic medical record of patient's work, which he has been promoting, since 2020.

IGNACIO DIEZ GONZALEZ has been a Cardiologist and the Head of the Chest Pain Unit of the Cardiology Service, Basurto University Hospital (Osakidetza-Basque Health Service), since December 2013, with specialization accreditation in heart disease diagnostic techniques. He is currently the Head of the BioCruces Clinical Cardiology Research Department. In relation to the subject of the study, he has participated as a Principal Investigator in the European Commission Project, Diagnostic Imaging Strategies for Patients With Stable Chest Pain and Intermediate Risk of Coronary Artery Disease (DISCHARGE), recently published in *The New England Journal of Medicine*.

BEGONYA GARCIA-ZAPIRAIN (Member, IEEE) was born in San Sebastián, Spain, in 1970. She received the degree in telecommunication engineering from the University of Basque Country, Spain, in 1994, and the Ph.D. degree in computer science and artificial intelligence from the University of Deusto, Spain, in 2004. From 2002 to 2008, she worked as the Director of the Telecommunication Department, University of Deusto, where she is currently a Full Professor. In 2001, she created the eVIDA Research Group, which is recognized by the Government of the Basque Country, Spain, and the European Network of Living Laboratories (ENoLL).