Investigation of Early-Stage Breast Cancer Detection using Quantum Neural Network

https://doi.org/10.3991/ijoe.v19i03.37573

Amjed Y. Sahib[1(✉)], Muazez Al Ali[2], Musaddiq Al Ali[3]
[1]College of Engineering, University of Wasit, Al Kut, Wasit, Iraq
[2]College of Dentistry, Al Ayen University, Nasiriyah, Iraq
[3]Department of Advanced Science and Technology, Toyota Technological Institute, Nagoya, Japan
ayousif@uowasit.edu.iq

Abstract—Aided image diagnostics (CAD) have been used in many fields of diagnostic medicine. It relies heavily on classical computer vision and artificial intelligence. Quantum neural network (QNN) has been introduced by many researchers around the world and presented recently by research corporations such as Microsoft, Google, and IBM. In this paper, the investigation of the validity of using the QNN algorithm for machine-based breast cancer detection was performed. To validate the learnability of the QNN, a series of learnability tests were performed alongside with classical convolutional neural network (CCNN). QNN is built using the Cirq library to perform the assimilation of quantum computation on classical computers. Series of investigations were performed to study the learnability characteristics of QNN and CCNN under the same computational conditions. The comparison was performed for real Mammogram data sets. The investigations showed success in terms of recognizing the data and training. Our work shows better performance of QNN in terms of successfully training and producing a valid model for smaller data set compared to CCNN.

Keywords—quantum neural network, breast cancer, classical neural network, machine learning, mammography

1 Introduction

With the advancement in medical and engineering fields, novel solutions were implemented to facilitate patients' health care and prolong their life [1–8]. The use of computer use of computer-aided diagnostic (CAD) is an important topic in engineering-medical research [9–10]. Recently, many researchers investigated the concept of automating CAD by building self-learning algorithms based on machine learning [11–13]. In order to build a successful diagnostic model of medical images, classical machine learning is used. However, training classical machine learning is consuming a huge computational resource in terms of the data set preparation as well as computer resources for the training phase [13–14]. For the medical diagnostic field, data are mostly visual-based, such as X-rays, computed tomography scan and magnetic

resonance imaging (MRI), etc. Therefore; computer vision tools are the most appropriate method to be used for CAD. Add to that, most data in modern diagnostic are computerized based [15]. As such, artificial intelligence is the most suitable to be used in the next generation of fully computerized CAD. Historically, CAD [16] has been approached in two sequential steps. The first step is to screen the data to detect the suspicious regions or what are so-called (Regions of interest) then, the region of interest will be "labeled" for the closest possible medical cases according to the corresponding disease likelihood. This is done by what so called expert systems [17–18]. Expert systems will assign diagnostic labels according to the probability of the region of interest diagnostic ascendingly, then eliminate medical cases of lower probabilities and then identify the highly likely disease. The problem with the use of expert systems is that they are built upon predetermined diagnostics for a fixed small number of cases and are mostly based on a wide variety of collective diagnostic data. For example, to diagnose a single case of apparent mass in appeared in mammography, a series of tests should be performed (Blood, tissue samples, multiple scans, etc.) to be fed to the expert systems than gives the preliminary diagnosis. On the other hand, artificial intelligence (AI) has great potential to replace traditional expert systems and allows one to reach a preliminary diagnosis in a very short time. AI is presented itself as a powerful tool for image classification and medical identification due to its characteristics of transforming representing information sets of data (database) into structured matrices of simple units containing weighted partial differential equations (PDEs). Weights then will be tuned to draw learning pathways as an intuitive mimic of the learning process of the biological neural system (i.e., the brain). The satisfying amount of learning data (i.e., raining data) to provide a robust outcome or valid model is depending on the AI model design and the problem itself. For data selection, two major aspects should be taken into consideration; the first one is identifying the objective of data to be trained for. In other words, what are the systematic methodologies of the doctor to make diagnostics based on such data. For example, in microscopic cultures, counting cells is one of the diagnostic tools; therefore, counting will necessitate isolating (through image segmentation) as well as labeling the targeted cells in each picture before submitting the data to the deep learning model. The second aspect is the data set size and number of items for each label. As it has been expressed before, training robustness is susceptible to data set size and the balance of data distribution in each label. The bottleneck of creating a strong and highly trusted CAD is the computational power [19–24]. For classical machine learning (which is performed on nowadays classical computers that use the binary system [0,1]), the researchers reach the computation limitations to modeling real-life problems such as biochemical interactions and immune system interaction with infection. Some algorithms are invented to rework the data feed and training process to optimize workflow for the available resources. Furthermore, some hardware-based methods are introduced to give the classical systems the needed dynamic storage to allocate the data feeds for the processing units as well as storing the tensor from the data (such as Intel Obtain-based modules and tensor processors). However, the largest computer (supercomputer) is still facing huge challenges to simulate the simulation tasks mentioned earlier. As such, in the medical field, classical computers are still facing serious limitations in terms of attaining good diagnoses compared to well-trained doctors with the same amount of data. To overcome classical computers' limitations, researchers are working

on developing so-called quantum computers. Quantum computers are applying the principles of entanglement and superpositions to the data unit, i.e., the bit to build the quantum bit. There are several approaches to physically implementing quantum computers such as photonic and silicon-based circuitry. However, the development is not moving fast as it has been anticipated due to entropy and noise problems. On the other hand, the logical aspects of a quantum computer are mature enough to be implemented as soon as a full-scale quantum computer is available. The advantages of a quantum computer can be shown by the phenomenal calculation speed and the data amount to be handled in a single calculation. The high speed and huge data processing ability of quantum computers are due to the nature of the method of conveying information itself, such that, instead of the binary representation of the data (i.e., classical bit), quantum bits within the quantum gates exchange information timelessly. The property of a quantum state makes one quantum computer hold calculation power equivalent to all existing classical computers (i.e., quantum supremacy) [25]. Quantum computers and computation are superior due to the nature of quantum information and quantum logic. Although the quantum computation field of study is relatively a new branch of applied mathematics, however; the physical and mathematical advancements of quantum computers are motivating the researchers to race toward implementing a unique type of logic gates and physical hardware. Nowadays, limited quantum bits computers are already serving in specific high technology applications such as pharmacy, and chemical interactions in next-generation batteries. It is anticipated that quantum computers will be available for public use with full capability at the end of the 21st century. The use of quantum computers for medical diagnostics can play a vital role in terms of the implementation of fast and systematic machine-based diagnostics of malignant tumors which are treatable if they can be diagnosed in the early stages. Breast cancer is the most frequent malignant tumor amongst women. It is a dominant cause of female mortality and is considered a serious public health problem all over the world. Current treatments for breast cancer include surgery, chemotherapy, immunotherapy, and radiation therapy. Breast cancer incidence and death rates increase with age but are mortality rate decrease significantly if it can be detected in the early stages, before the metastasizing phase. The eradication and therapeutic success of breast cancer are related to tumor stratification and dissemination. Breast tumors whether they are benign or malignant are distinguished into four major classes, based on size, age, node involvement, and tumor grade. These stages are 1; consists of the well-defined and localized tumor mass, characterized by poor invasion properties. Stage 2 and 3, corresponds to an increased tumor volume and acquisition of invasive phenotype. The metastasis dissemination and huge tumor size with invasive phenotype are classified as stage 4. Chemotherapy, radiation, and targeted therapies have made major advances in patient management over the past decades, but refractory diseases and recurrence remain common. The early-stage diagnostic will lead to treating cancer before the metastasis stage, at which cancer will attack different organs by migrating through lymph nodes. A mammogram is one of the popular tests for breast cancer early detection. Mammograms have been used efficiently to reduce the mortality rate of women with breast cancer. Early detection is based on the oncologist's exam of the x-ray image and then examining the suspicious tissue by taking a biopsy. However, due to the limitation of highly trained medical staff, cancer false positive is common in mammogram image detection

as well as false negative [26–28] due to the misconception of less-skilled eyes for the masses in the image. Another point worth mentioning is that mammogram is an X-ray image that take short time to be made, yet the speed of mass examination is related to how many doctors exist and their level of experience. Therefore, CAD is a promotion point as a necessary way to reduce diagnostic time [29–31] and decrease cost per test by reducing the number of specialists in the mammogram testing unit in hospitals (This reduction will lead to remobilizing the manpower to other sections within the hospital which will lead to double the efficiency of the hospital's workforce) as well as the reduction of the number of false-positive by double-checking the X-ray image by the doctor and the computer. This paper is examining the use of QNN to build an "ultra-intelligent machine". To attain this goal, the following Key questions should be answered: Firstly, how to build a QNN and how to transform the classical data from plane mammogram images to become quantum data. The second key question is how to evaluate the learnability of QNN. The third question is, how is implementing a benchmarking of learnability evaluation of QNN? The final question is: can we deduce that, QNN is able to perform successful mammography diagnostic in the future according to the current investigation? As such, this study introduced the learnability factors. Furthermore, a series of numerical investigations were conducted to examine the various aspects that govern AI with QNN. Moreover, a hybrid model of CCNN and QNN is introduced to allow the implementation of quantum logic in the near future instead of waiting for a fully capable quantum computer to conduct breast cancer early detection for mammography. The layout of the paper is the following. Section 2 deals with the learning of classical and quantum neural networks. Whereas section 3 discusses the layout of quantum and classical neural networks and mammogram data structure. In section 4, we will present the results and discuss the outcome. Finally, the conclusions are presented in section 5.

2 Classical and quantum neural network

The core of a successfully implemented machine learning system is simple. Information nowadays is not just conception only; it is a physical entity too. Nature around us gathers information at the subatomic level, in which the line between information and physical matter can interchange, extrapolate, and convert. Electronic communication technologies, from Samuel's revolutionary bubbles machine in 1809 to the fast fiber optics connections, rely on summarizing the information in simple possible unit codes from codes. Revolutionized computation was achieved by adopting a binary state (0 and 1) to represent information. The unit binary code necessitates special handling to produce useful information, such that the interpretation and processing will need a logic unit designated to the binary process. The binary system is one of the most reliable sources of information transition recently. Despite the advancement of modern classical computers and information transfer infrastructure, the classical computers established based on a binary system are limited compared to the complex nature of real-world information exchange and physical problems. To be solved on classical computers, simulations and other issues needed first to be simplified to a certain degree to be solvable with maintaining a reasonable degree of accuracy. This limitation is intriguing for the researcher to revolutionize the methodologies to approach the informatics

of real-world problems. The advancement of physics has introduced more sophisticated ways to exchange information, i.e., quantum informatics. The quantum state is a semi-timeless way of exchanging information [32–34]. Driven by the power of information, it was a belief in the 19th century that the physical state could be changed without energy loss by changing the information of the subatomic system and controlling the disorder to make order (equilibrium of the physical system) [35]. The information in the quantum world has unique properties that need a different approach than classical information logic. Machine learning is the transcending form of the neural network, which was immersed first in the form of a "logic theorist program" that was invented in 1956 [36]. The neural network process generally starts by discretizing the problem (X) (as shown in Figure 1) to be distributed as simple subsystems, then fitting each subsystem to arbitrary function ψ. Functions output will be submitted to a comparator Ω which judges the data based on its probability function the gives the generalized output of the subs ystems Ψ.

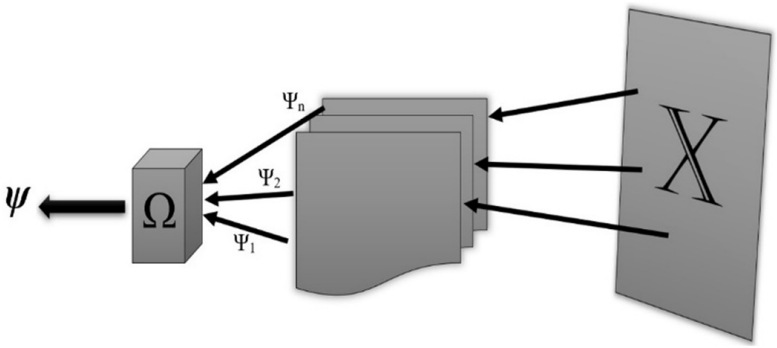

Fig. 1. Machine intelligence process unit

Machine learning is a network of node clusters; each node has a weighted sub-function to be tuned. This tuneable sub-function is representative of a fraction of the analyzer model. The node's output mimics neural cells by adopting an activation function to control the output with predefined criteria. Activation function can take the form of rectified linear, Sigmoid, or hyperbolic. These nodes called neurons are clusters into stacks called layers which can be dense layers or unique function layers such as convolution layers. Training of the neural network is performed by tuning each neuron's weights for the optimum value that fits the training data. Here, optimization algorithm selection plays a significant role in training success. With increasing the training rate, data loss for each prediction will drop. If the data is insufficient or the neural network is designed poorly, the fit convergence cannot reach sufficient value. In this case, the under-fitting problem is occurring. Contrary to the under-fitting problem, the overfitting problem occurs when the convergence is satisfied for the model; however, the neural network has the poor capability to predict and recognize input data that is not in the training set. Quantum machine learning is a subject of the quantum computation branch that has been gaining attention in recent decades, which emerged from quantum computations. Quantum computation has emerged from the statistical implementation of the work of quantum theory as the work of Helstorm [37] and Holevo [38]. Quantum computer concepts were researched by many researchers, such as Benioff [39] and Fynman [40–41].

Quantum computers, due to their design, have the upper hand in extensive data analysis [42]. Quantum computers unit information, i.e., a quantum bit (Qbit), shares the same characteristics of quantum concepts, such that it is in a simple state of a linear superposition of 0 and 1 [43–44]. To understand the behavior of computation in the quantum neural network, we will introduce a basic mathematical notation that is commonly used to express computation in a quantum environment. In other words, Dirac's bra (⟨|) ket notation is used to represent quantum states as vectors, where |⟩ and ⟨| represent a column vector and the conjugate transpose of a vector, respectively as shown in Figure 2.

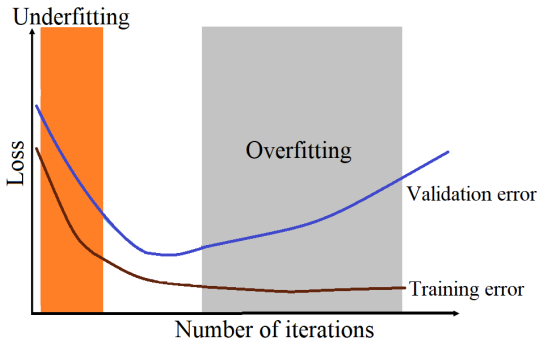

Fig. 2. Under and overfitting problem of neural network

Accordingly, the fundamental quantum states $|0\rangle$ and $|1\rangle$ can be expressed in Dirac's notation respectively as:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1}$$

The standard quantum operator representation of Pauli matrices can be defined as:

$$\widehat{\sigma_1} = |0\rangle\langle 1| + |1\rangle\langle 0| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2}$$

$$\widehat{\sigma_2} = -i|0\rangle\langle 1| + i|1\rangle\langle 0| = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{3}$$

$$\widehat{\sigma_3} = |0\rangle\langle 0| - i|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4}$$

which are Hermitian and unitary. The unitary operator on the two-dimensional Hilbert space, spanned by the basis $\{|0\rangle, |1\rangle\}$ is defined by $\hat{1} = |0\rangle\langle 0| + |1\rangle\langle 1|$, which takes the form of an identity matrix(I). Now we have the representation of unitinformation within the quantum computers, so special logic operator controllers i.e., quantum

gates, are needed to interpret and operate a program using this quantum unit. Quantum gates are represented mathematically with a set of matrices representing the probability of a unified state within the estate. The Walsh-Hadamard transform unitary operator is an Example of quantum gates that map the basic unites to superposition state, given by:

$$\widehat{U}_{WH} = \frac{\widehat{\sigma_1} + \widehat{\sigma_3}}{\sqrt{2}} \tag{5}$$

In addition to the above definition, we define the binary alphabet by $A_2 = \{0, 1\}$, and for a set of L-length binary strings greater than 1, we denoted it by A_2^L. After these basic definitions, we now turn to address some of the fundamental quantum computation properties in the neural network. A quantum neuron is a basic unit in a quantum neural network (QNN). A neural firing operator that describes the firing of a two-level neuron [29–30] is given by:

$$\widehat{F} = \frac{\hat{1} - \widehat{\sigma_3}}{2} F \tag{6}$$

Where F is the firing frequency of neurons in Hertz, the eigenvectors of this operator are given by :

$$\widehat{F}\left|s\right\rangle = sF\left|s\right\rangle, s = 0,1 \tag{7}$$

Therefore, these two eigenvectors $|0\rangle$ and $|1\rangle$ correspond to neural activity where firing frequency 0 Hz and 1 Hz, respectively. In other words, two quantized energy levels can be obtained from each neuron Hamiltonian and can be expressed in Eq. 9 [29–30]:

$$\widehat{H} = 2\pi\eta\widehat{F} \tag{8}$$

Consequently, the energy levels associated with eigenvectors $|0\rangle$ and $2\eta F$, respectively. In terms of neural networks with L-neurons, Eq. 7 can be generalized to be:

$$\widehat{F}_L\left|s_1 s_2 \ldots s_L\right\rangle = s_L F\left|s_1 s_2 \ldots s_L\right\rangle \tag{9}$$

Where $\widehat{F_L}$ the firing frequency field operators of Lth neurons and any pair of neural firing operators commute; that is for $L, K = 1, 2 \ldots d, \left[\widehat{F_L}, \widehat{F_K}\right] = 0$. Thus, the total neural firing frequency operator is given by:

$$\widehat{F}_{Tot} = \sum_{L=1}^{d} \widehat{F}_L \tag{10}$$

In terms of eigenvalue for neural network (10) can be written as:

$$\widehat{F}_{Tot}\left|s_1 s_2 \ldots s_L\right\rangle = \sum_{L=1}^{d} s_L F\left|s_1 s_2 \ldots s_L\right\rangle \tag{11}$$

The quantum circuits implemented a dynamics L-neuron QNN is a chain of unitary operators. The exception is that the conditional unitary operator implemented for the operator follows the link of neural structure. Only conditional operators respect

network processing direction and topology can be implemented in QNN. Mathematically, we can express an N-length quantum computational that propagates forward in a chain comprised of a sequential product of unitary operators as:

$$\hat{O} = \hat{U}_N \hat{U}_{N-1} \ldots \hat{U}_1 \tag{12}$$

The chain in Eq. 12 represents the unitary forward operation, where it is applied such that U_1 is the first operator and U_L is the final. The backward propagation of the network is the conjugate transpose of Eq. 12 given by:

$$\hat{O} = \hat{U}_1^{\dagger} \ldots \hat{U}_{N-1}^{\dagger} U_N^{\dagger} \tag{13}$$

Formally, given the general initial quantum neural field dynamics density of QNN $\hat{\rho}_{in} = \Sigma_{r,s \in A_2^d} \rho_{r,s|r\rangle\langle s|}$, the propagation of the quantum network system can be written as:

$$\hat{\rho}_{out} = \hat{O}\hat{\rho}\hat{O}^{\dagger} = \sum_{r',s' \in A_2^d} \left(\begin{array}{l} \sum_{r,s \in A_2^d} \rho_{r,s} \left\langle r' \middle| \hat{U}_1 \ldots \hat{U}_{N-1} \hat{U}_N \middle| r \right\rangle \\ \left| r' \right\rangle \left\langle s' \right| \left\langle s \middle| \hat{U}_1^{\dagger} \ldots \hat{U}_{N-1}^{\dagger} \hat{U}_N^{\dagger} \middle| s' \right\rangle \end{array} \right) \tag{14}$$

Where r, s represents the input firing patterns and r', s' the output patterns for the neurons, respectively. Moreover, the computation process propagates backward and forward of the chain, such that the propagation in the forward direction from begging to the end system corresponds to the amplitude $\langle r' | \hat{U}_1 \ldots \hat{U}_{N-1} \hat{U}_N | r \rangle$ and $\langle s | \hat{U}_1^{\dagger} \ldots \hat{U}_{N-1}^{\dagger} \hat{U}_N^{\dagger} | s' \rangle$ represents backward prorogation amplitude. When $r' \neq s'$ the system is unbalanced and doesn't reach the final solution, this is called mismatching between output firing dynamics. In other words, the computation amplitude $\langle r' | \hat{U}_1 \ldots \hat{U}_{N-1} \hat{U}_N | r \rangle$ in the forward direction does not match the computation amplitude $\langle s | \hat{U}_1^{\dagger} \ldots \hat{U}_{N-1}^{\dagger} \hat{U}_N^{\dagger} | s' \rangle$ in the backward direction, which violates computation in quantum respective. However, when $r' = s'$ this means the computed output in both directions are matched, and the final density operator is a diagonal component for each $s' \in A_2^d$ can be expressed as:

$$\left\langle s' \middle| \hat{\rho}_{out} \middle| s' \right\rangle \middle| s' \right\rangle \left\langle s' \right| =$$

$$\left(\sum_{r,s \in A_2^d} \rho_{r,s} \left\langle r' \middle| \hat{U}_1^{\dagger} \ldots \hat{U}_{N-1}^{\dagger} \hat{U}_N^{\dagger} \middle| r \right\rangle \middle| r' \right\rangle \left\langle s' \right| \left\langle s \middle| \hat{U}_1^{\dagger} \ldots \hat{U}_{N-1}^{\dagger} \hat{U}_N^{\dagger} \middle| s' \right\rangle \right) \middle| s' \right\rangle \left\langle s' \right| \tag{15}$$

Eq.15 tells us that the final output firing pattern for a neuron $s' \in A_2^d$ is the weighted sum average of the initial pair firing pattern propagated in forward and backward of the computational system chain.

In the computer context, this both-direction propagation is the primary result of QNN based on a quantum circuit (which can be generalized to any quantum system exhibiting both-direction propagation). In a computer since perspective, the propagation forward and backward can be explained as dynamics prob and response, respectively; returning to Eq. 14 the term $\langle r' | \widehat{U}_1 \ldots \widehat{U}_{N-1} \widehat{U}_N | r \rangle$ represents the response from the beginning to the end of the system computation; while the term $\langle s | \widehat{U}_1^{\dagger} \ldots \widehat{U}_{N-1}^{\dagger} \widehat{U}_N^{\dagger} | s' \rangle$ represents the response from the beginning to the end of the system computation; where output firing s' and input firing s are linked. As we mentioned above, when $s' \neq r'$ a mismatch between prob and response occurs, and there is no solution can be found; while when $s' = r'$ this means the prob and response are matched, which produces an echo with intensity given in Eq. 15. The computation, in this case, can be described as searching for a solution for a computational system, where the alternative output for specific prob results in specific intensity as a response. This process occurs simultaneously; the computation between both directions is simultaneous to reach the final solution. The mentioned fundamental characterizes the quantum computation system and is not limited to the QNNS process.

3 Materials and methods

Mammogram image data set is provided by [45] for 6 breast abnormality classes: calcification, circumscribed masses, spiculate masses, architectural distortion, asymmetry, and healthy breast images. The data set consists of the training set of 45000 images for all the previously mentioned 6 classes and 7500 images as the validation set. Each image is of size 1080 by 1080 pixels. MATLAB was implemented to investigate CNN design and perform training. Because quantum-based libraries provided by quantum computer service providers (such as IBM and GOOGLE) are built with Python, the Python version of the optimal MATLAB CNN was used as a baseline comparison program with QNN. Quantum gates formation and QNN were built using python language. For classical coding, CNN was built with multi-stages of convolutional layers to compress the data stream to the throttle point (as shown in the machine learning architecture in Figure 3). Images are needed to be resized to decrease the pixel numbers to smooth the training; add to that, reducing pixel numbers is necessary to perform training with reasonable computer resources. Using raw images of high resolution will necessitate an increase in the node's number inside the neural network, which will not necessarily improve the model recognition ability.

As such, in our research, input images were compressed to 256 by 256 pixels and uploaded using the computer unified device architecture's library (CUDA) to enable the parallelization of tensor multiplication. Several CNN models were built to adjust the optimal neuron number, starting from 50 million neurons to 2 million neurons.

The best CNN design and the optimal number of neurons are evaluated based on the model's learnability and validation. Learnability is the most important aspect of machine learning, such that it will show if the program or the set is in fact, machine intelligence

or not. The learnability can be evaluated by the behavior of the machine learning to data and the change of its architecture in a way that it should always show the symptoms of acquiring the information. One of the easiest symptoms to be measured is the under and overfitting of the machine model to the training data. For example, a neural network by design might not converge sufficiently to give a valid model. This phenomenon can be referred to as underfitting. Underfitting can happen for many reasons, such as poor design models or insufficient data. On the other hand, if the learning rate is converged too much, the model recognizes the training set flawlessly. Yet, it lacks the ability to recognize the new images that the model is never experienced in the training phase. This phenomenon is called overfitting. Overfitting can happen due to the lack of trainable neurons in each training cycle. The transformation from underfitting to overfitting of the neural network strongly indicates the successful learnability of the implemented network design. On the other hand, our QNN was built using the Python language framework associated with the Cirq library as a quantum circuit framework [44–48]. Cirq library allows to perform of quantum computation on classical computers by assimilating the Qbit by its representative classical bits. This is the only way to test full-scale quantum programming due to the lack of usable quantum computers at the current time. However, due to the nature of Qbit, a huge amount of storage is needed due to the complex nature of Qbit and its phenomenal information capacity. As such, performing full-scale, and high or even moderate resolution image training with QNN on classical computers is extremely difficult due to the lack of computational resources to perform superposition, therefore; images must be sufficiently compressed to allow the computer to perform a quantum-based search algorithm. In this paper, QNN work frame is as follows: The first step is to load the images into the model and store the data in the RAM as tensors of classical bits. Because quantum computation handles only quantum data, it is inevitable to transform classical bits into Qbit using the Cirq library, which is the second step of the QNN work-frame. Transformation is performed by mapping bits into a Qbit matrix using the transform function based on tensor products of Bits [49–54]. The third step is to build the quantum circuit of sequential quantum gate formations, as shown in Figure 4. In this paper, a quantum circuit consists of a series of Ising gates that show good recognition ability. Starting from 50 x 50 pixels image compression, QNN simulation consumed computer resources before training started. In this research for our QNN design investigation, our model could not work for the full training and validation sets unless the image compression was extremely to 4 x 4 pixels (as presented in Figure 5). After compression, the classical bits are transformed to the Qbit using the previously mentioned method. After wrapping up the compressed training data set in terms of the quantum state (i.e., Qbit), the training of the qubit will start. The final training matrix will be exported and saved to be implemented in the prediction program. Our QNN algorithm is summarized in Figure 6.

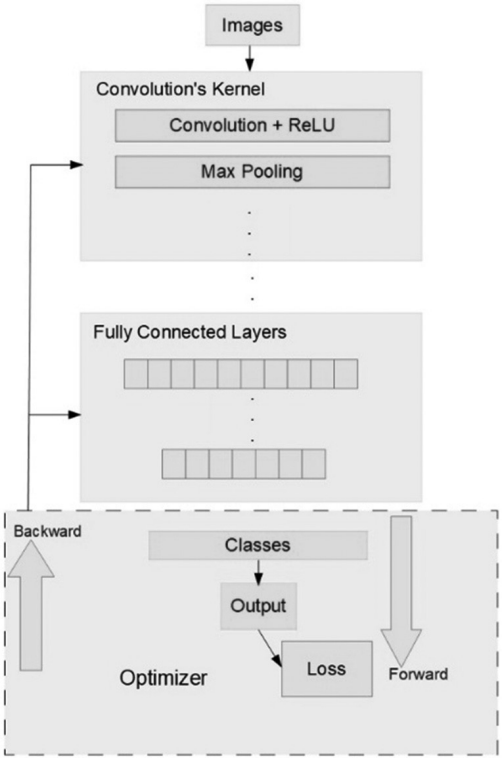

Fig. 3. Classical convolution neural network architecture

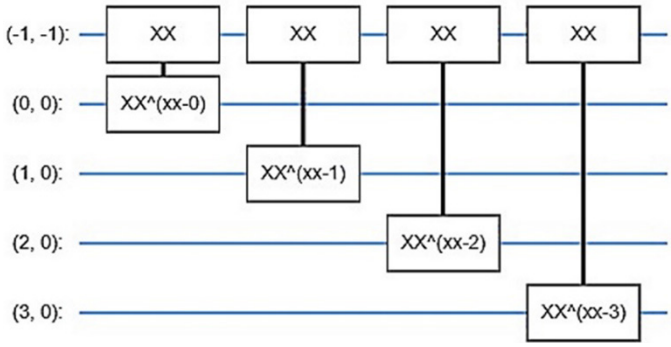

Fig. 4. Ising quantum computation circuit for mammogram image-based cancer detection

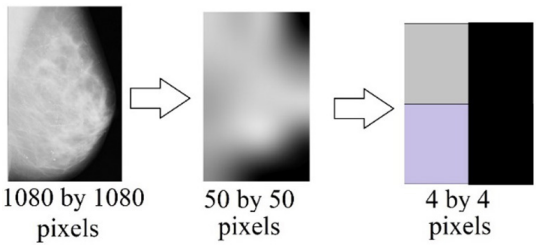

Fig. 5. Image compression representation before full-scale QNN training

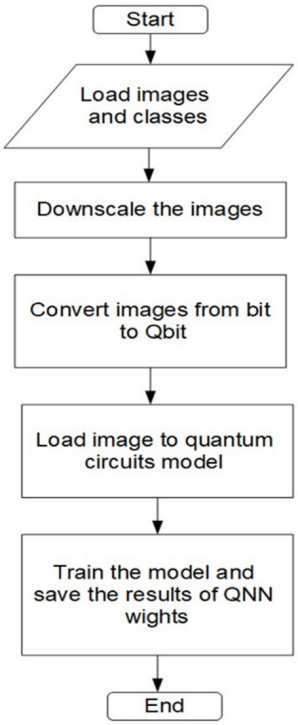

Fig. 6. QNN identification algorithm

4 Results and discussion

In this section we are showing the results of our investigation of the use of CCNN and QNN

4.1 The investigation of CCNN models

As been previously explained in the material and method section, CCNN was examined through several designs by increasing the neuron count from 2 million to

50 million. A high number of neurons is needed, so mammogram image classification is difficult even for human eyes. On 50 million neurons, the model show overfitting, as shown in Figure 7.

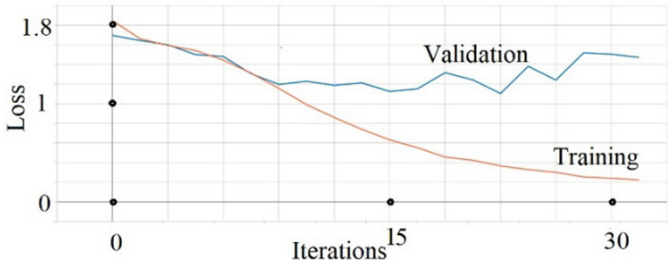

Fig. 7. Classical neural network overfitting

While with 2 million neurons, the model shows underfitting behaviour (as shown in Figure 8).

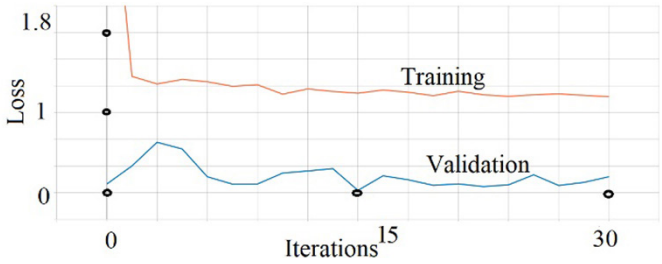

Fig. 8. Classical neural network underfitting

Such a shift from underfitting to overfitting proves that the CCNN has a successfully learnability [55]. As shown in Figure 9, the best natural network was the 1.4 million neurons model, while 50 million nodes now showed more than 53% accuracy (Figure 10).

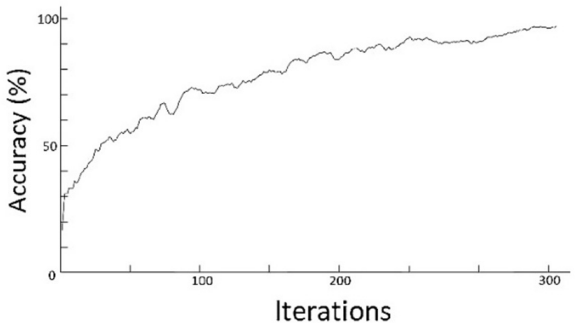

Fig. 9. Accuracy history of the best classical neural network model

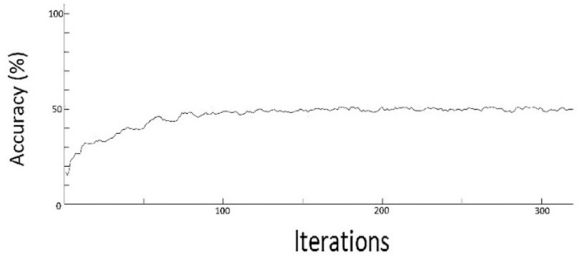

Fig. 10. Accuracy history of the under-fitted classical neural network model

4.2 The investigation of QNN models

The capacity to learn is the most crucial component that measures the feasibility of using QNN as a mammogram diagnostic tool. The machine learning's response to input and the modification of its design in such a manner that it consistently displays the signs of learning may be used to assess the learnability. The under- and overfitting of the machine model to the training data are used as intuitive signs to validate the QNN learnability. As such, using the same analogy to examine QNN as an image classifier, a model showed a transformation from underfitting (Figure 11) to overfitting (Figure 12). As mentioned earlier, because QNN was simulated on a classical computer, the image must be highly compressed to give the required RAM and CPU resources for constructing a Qbit. Although the mammogram images were highly compressed, QNN could reach converged after 5 training trials to achieve 58% accuracy (Figure 13) with stead learning speed. Even though the training of highly compressed images, QNN showed sign of learnability by transforming the loss from underfitting to overfitting. This sign has proven the ability to use a large model of the QNN on a full-scale quantum computer to produce a robust CAD model. Also, QNN showed a high tendency to achieve a high fitting level with just several training cycles. This is an anticipated result due to the nature of the quantum gate operations.

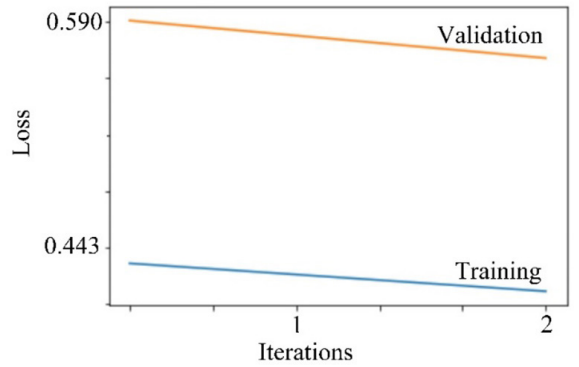

Fig. 11. Quantum neural network underfitting

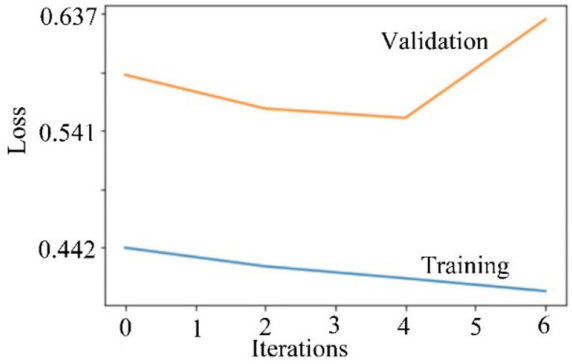

Fig. 12. Quantum neural network overfitting

4.3 The investigation of CCNN and QNN models using same data set

To verify the learning ability of quantum and classical models in fair terms (i.e., examine both models under similar conditions), in this section we performed the CCNN and QNN evaluation for the training and validation data sets that have similar data compression and for the same number of samples for both models.

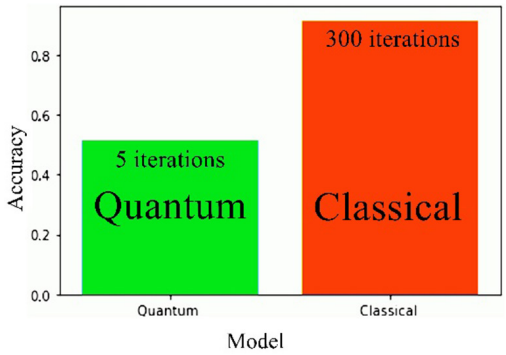

Fig. 13. Classical and quantum network predictability

Because it is impossible to train both QNN and CNN using the 45000 images with reasonable image compression, down-sampling of the full-scale data was examined to show which model is faster in achieving stable accuracy growth. Two data sets were chosen for this study. The first set has 600 training images, and the second set of 1200 images. Both training cases have a validation set of 300 images. This choice was selected to ensure CCNN's fair learning chance and QNN stability during testing, so if the number of samples is chosen over 2000, it will lead to the overload of the computer memory. As such, for the data set of 600 images for the training set and 300 images for the validation test are compressed to 50 x 50 pixels (as shown in Figure 14), and used for both CCNN and QNN models.

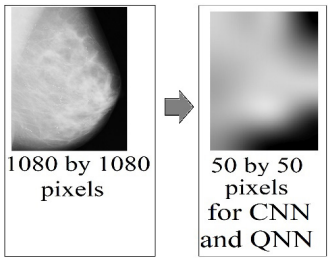

Fig. 14. Image compression for the new data set

Training for this data set showed a remarkable convergence time of QNN compared to CCNN (as shown in Figure 15). Validation accuracy for QNN showed superior performance compared with CCNN. CCNN showed fluctuation in validation accuracy, which can be interpreted as underfitting for the classical model. By increasing the training set to 1000 images to decrease the classical model's underfitting; QNN maintains its superiority compared to CCNN, as shown in Figure 16.

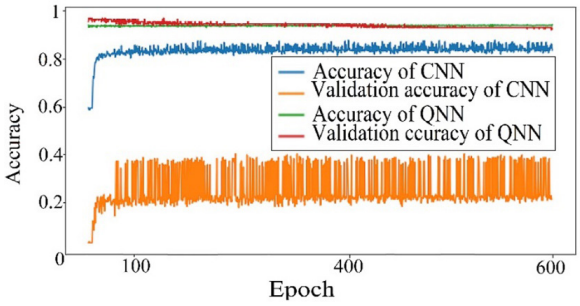

Fig. 15. Classical and quantum network accuracy for 600 images data set

Mathematically speaking, the ability to parallelize the calculation process within the Qbit will lead to decreasing the nonlinearity due to continuous bidirectional feedback of neurons, as well as compressing the calculation process with maintaining robustness. For the foregoing reasons, QNN can achieve rapid, stable accuracy with a shorter time and fewer samples than CCNN.

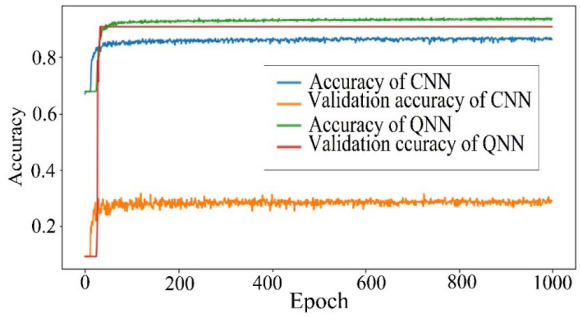

Fig. 16. Classical and quantum network accuracy for 1000 images data set

It is valid to say that QNN can give a reasonable prediction model for a much smaller data set than CNN. Applying QNN on a fully capable quantum computer will provide faster training with smaller data set. Under such circumstances, QNN will provide fast mass medical imaging diagnostic with a minimum false negative.

5 Conclusion

In this research, we performed an investigation of early breast cancer detection using quantum computation by training mammogram images to a designed QNN model. The QNN model was successfully used to train actual mammogram data sets for comparison. The investigation demonstrated effectiveness in terms of training and data recognition. In comparison to CCNN as a mammogram image diagnostic, our study demonstrates the advantage of QNN in effectively training and delivering a viable model for less time and high accuracy. From this investigation, we conclude that the anticipated availability of real and robust quantum computers for commercial use will be a strong asset for robust mass mammogram detection successfully for a significantly short time with high accuracy. In other words, the main question discussed in this paper is whether it is possible to implement in theory, specific real data of early cancer detection in a quantum-based has been answered positively. It can be said that this objective was achieved by satisfying the criteria of learning for the designed QNN model. In addition, we conclude that QNN may provide a credible prediction model with a significantly smaller data set, unlike CCNN, which makes QNN an anticipated tool for performing mass medical imaging diagnostics with a low false-negative rate.

5.1 Data availability

All results in this paper are calculated by using in-house MATLAB and Python codes. The code cannot be shared this time as it is used in ongoing work. However, all results can be reproduced by adopting the same assumptions.

6 References

[1] M. Learning, "Heart disease diagnosis and prediction using machine learning and data mining techniques: a review," Advances in Computational Sciences and Technology, vol. 10, no. 7, pp. 2137–2159, 2017.

[2] L. P. Dasi, H. A. Simon, P. Sucosky, and A. P. Yoganathan, "Fluid mechanics of artificial heart valves," Clinical and Experimental Pharmacology and Physiology, vol. 36, no. 2, pp. 225–237, 2009. https://doi.org/10.1111/j.1440-1681.2008.05099.x

[3] E. Cuadrado-Godia et al., "Cerebral small vessel disease: a review focusing on pathophysiology, biomarkers, and machine learning strategies," Journal of Stroke, vol. 20, no. 3, p. 302, 2018. https://doi.org/10.5853/jos.2017.02922

[4] M. Al Ali, A. Y. Sahib, and M. Al Ali, "Teeth implant design using weighted sum multi-objective function for topology optimization and real coding genetic algorithm," in The 6th IIAE International Conference on Industrial Application Engineering 2018, 2018, pp. 182–188. https://doi.org/10.12792/iciae2018.037

[5] M. Al Ali, A. Y. Sahib, and M. Al Ali, "Design light weight emergency cot with enhanced spinal immobilization capability," in 6th Asian/Australian Rotorcraft Forum & Heli Japan, 2017, pp. 1–11.
[6] M. Al Ali, M. Al Ali, R. S. Saleh, and A. Y. Sahib, "Fatigue life extending for temporomandibular plate using non parametric cascade optimization," in Proceedings of the World Congress on Engineering 2019, 2019, pp. 547–553.
[7] M. A. Al-Ali, M. A. Al-Ali, A. Takezawa, and M. Kitamura, "Topology optimization and fatigue analysis of temporomandibular joint prosthesis," World Journal of Mechanics, vol. 7, no. 12, pp. 323–339, 2017. https://doi.org/10.4236/wjm.2017.712025
[8] R. S. Abass, M. Al Ali, and M. Al Ali, "Shape and topology optimization design for total hip joint implant," in Proceedings of the World Congress on Engineering, 2019, vol. 0958.
[9] L. Heflin, S. Walsh, and M. Bagajewicz, "Design of medical diagnostics products: a case-study of a saliva diagnostics kit," Computers & Chemical Engineering, vol. 33, no. 5, pp. 1067–1076, 2009. https://doi.org/10.1016/j.compchemeng.2008.09.024
[10] H. E. Pople, "Heuristic methods for imposing structure on ill-structured problems: the structuring of medical diagnostics," Artificial Intelligence in Medicine, vol. 51, pp. 119–190, 1982. https://doi.org/10.4324/9780429052071-5
[11] J. G. Richens, C. M. Lee, and S. Johri, "Improving the accuracy of medical diagnosis with causal machine learning," Nature Communications, vol. 11, no. 1, pp. 1–9, 2020. https://doi.org/10.1038/s41467-020-17419-7
[12] N. A. Valentine, T. M. Alhawassi, G. W. Roberts, P. P. Vora, S. N. Stranks, and M. P. Doogue, "Detecting undiagnosed diabetes using glycated haemoglobin: an automated screening test in hospitalised patients," Medical Journal of Australia, vol. 194, no. 4, pp. 160–164, 2011. https://doi.org/10.5694/j.1326-5377.2011.tb03762.x
[13] C. Jacobs and B. van Ginneken, "Google's lung cancer AI: a promising tool that needs further validation," Nature Reviews Clinical Oncology, vol. 16, no. 9, pp. 532–533, 2019. https://doi.org/10.1038/s41571-019-0248-7
[14] S. Kulkarni, N. Seneviratne, M. S. Baig, and A. H. A. Khan, "Artificial intelligence in medicine: where are we now?," Academic Radiology, vol. 27, no. 1, pp. 62–70, 2020. https://doi.org/10.1016/j.acra.2019.10.001
[15] A. Y. Sahib, H. Seyedarabi, R. Afrouzian, and M. Farhoudi, "A MATLAB-based toolbox to simulate transcranial direct-current stimulation using flexible, fast, and high quality tetrahedral mesh generation," IEEE Access, vol. 10, pp. 76573–76585, 2022. https://doi.org/10.1109/ACCESS.2022.3190410
[16] K. Doi, "Current status and future potential of computer-aided diagnosis in medical imaging," The British Journal of Radiology, vol. 78, no. suppl_1, pp. s3–s19, 2005. https://doi.org/10.1259/bjr/82933343
[17] S. A. Naser, R. Al-Dahdooh, A. Mushtaha, and M. El-Naffar, "Knowledge management in ESMDA: expert system for medical diagnostic assistance," AIML Journal, vol. 10, no. 1, pp. 31–40, 2010.
[18] K.-P. Adlassnig, "A fuzzy logical model of computer-assisted medical diagnosis," Methods of Information in Medicine, vol. 19, no. 03, pp. 141–148, 1980. https://doi.org/10.1055/s-0038-1636674
[19] M. Al Ali, A. Takezawa, and M. Kitamura, "Comparative study of stress minimization using topology optimization and morphing based shape optimization comparative study of stress minimization using topology optimization and morphing based shape optimization," 2018.
[20] M. Al Ali and M. Shimoda, "Toward multiphysics multiscale concurrent topology optimization for lightweight structures with high heat conductivity and high stiffness using MATLAB," Structural and Multidisciplinary Optimization, vol. 65, no. 7, pp. 1–26, 2022. https://doi.org/10.1007/s00158-022-03291-0

[21] M. Al Ali and M. Shimoda, "Investigation of concurrent multiscale topology optimization for designing lightweight macrostructure with high thermal conductivity," International Journal of Thermal Sciences, vol. 179, p. 107653, 2022. https://doi.org/10.1016/j.ijthermalsci.2022.107653

[22] M. Fujioka, M. Shimoda, and M. Al Ali, "Concurrent shape optimization for multiscale structure with desired static deformation," The Proceedings of The Computational Mechanics Conference, vol. 2021.34, p. 3, 2021. https://doi.org/10.1299/jsmecmd.2021.34.003

[23] M. Fujioka, M. Shimoda, and M. Al Ali, "Shape optimization of periodic-microstructures for stiffness maximization of a macrostructure," Composite Structures, vol. 268, p. 113873, 2021. https://doi.org/10.1016/j.compstruct.2021.113873

[24] M. Fujioka, M. Shimoda, and M. Al Ali, "Concurrent shape optimization of a multiscale structure for controlling macrostructural stiffness," Structural and Multidisciplinary Optimization, vol. 65, no. 7, pp. 1–27, 2022. https://doi.org/10.1007/s00158-022-03304-y

[25] F. Arute et al., "Quantum supremacy using a programmable superconducting processor," Nature, vol. 574, no. 7779, pp. 505–510, 2019.

[26] P. T. Huynh, A. M. Jarolimek, and S. Daye, "The false-negative mammogram," Radiographics, vol. 18, no. 5, pp. 1137–1154, 1998. https://doi.org/10.1148/radiographics.18.5.9747612

[27] W. Spiesberger, "Mammogram inspection by computer," IEEE Transactions on Biomedical Engineering, no. 4, pp. 213–219, 1979. https://doi.org/10.1109/TBME.1979.326560

[28] I. T. Gram, E. Lund, and S. E. Slenker, "Quality of life following a false positive mammogram," British Journal of Cancer, vol. 62, no. 6, pp. 1018–1022, 1990. https://doi.org/10.1038/bjc.1990.430

[29] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," Computerized Medical Imaging and Graphics, vol. 31, no. 4–5, pp. 198–211, 2007. https://doi.org/10.1016/j.compmedimag.2007.02.002

[30] R. M. Nishikawa, "Current status and future directions of computer-aided diagnosis in mammography," Computerized Medical Imaging and Graphics, vol. 31, no. 4–5, pp. 224–235, 2007. https://doi.org/10.1016/j.compmedimag.2007.02.009

[31] K. Doi, H. MacMahon, S. Katsuragawa, R. M. Nishikawa, and Y. Jiang, "Computer-aided diagnosis in radiology: potential and pitfalls," European Journal of Radiology, vol. 31, no. 2, pp. 97–109, 1999. https://doi.org/10.1016/S0720-048X(99)00016-9

[32] P. Zanardi, "Quantum entanglement in fermionic lattices," Physical Review A, vol. 65, no. 4, p. 42101, 2002. https://doi.org/10.1103/PhysRevA.65.042101

[33] G. Ghirardi, L. Marinatto, and T. Weber, "Entanglement and properties of composite quantum systems: a conceptual and mathematical analysis," Journal of Statistical Physics, vol. 108, no. 1, pp. 49–122, 2002. https://doi.org/10.1023/A:1015439502289

[34] K. A. Dennison and W. K. Wootters, "Entanglement sharing among quantum particles with more than two orthogonal states," Physical Review A, vol. 65, no. 1, p. 10301, 2001. https://doi.org/10.1103/PhysRevA.65.010301

[35] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, "Thermodynamics of a physical model implementing a Maxwell demon," Physical Review Letters, vol. 110, no. 4, p. 40601, 2013. https://doi.org/10.1103/PhysRevLett.110.040601

[36] R. Millstein, "The logic theorist in LISP," International Journal of Computer Mathematics, vol. 2, no. 1–4, pp. 111–122, 1968. https://doi.org/10.1080/00207166808803027

[37] C. W. Helstrom, "Quantum detection and estimation theory," Journal of Statistical Physics, vol. 1, no. 2, pp. 231–252, 1969. https://doi.org/10.1007/BF01007479

[38] A. S. Holevo, "Testing statistical hypotheses in quantum theory," Probability and Mathematical Statistics, vol. 3, pp. 113–126, 1982.

[39] P. Benioff, "Quantum mechanical Hamiltonian models of turing machines," Journal of Statistical Physics, vol. 29, no. 3, pp. 515–546, 1982. https://doi.org/10.1007/BF01342185
[40] R. P. Feynman and others, "Simulating physics with computers," Int. J. Theor. Phys, vol. 21, no. 6/7, 1982. https://doi.org/10.1007/BF02650179
[41] R. P. Feynman, "Quantum mechanical computers," Optics News, vol. 11, no. 2, pp. 11–20, 1985. https://doi.org/10.1364/ON.11.2.000011
[42] A. Ekert and R. Jozsa, "Quantum computation and Shor's factoring algorithm," Reviews of Modern Physics, vol. 68, no. 3, p. 733, 1996. https://doi.org/10.1103/RevModPhys.68.733
[43] A. Y. Kitaev, A. Shen, M. N. Vyalyi, and M. N. Vyalyi, "Classical and quantum computation," no. 47. American Mathematical Soc., 2002. https://doi.org/10.1090/gsm/047
[44] D. P. DiVincenzo, "Quantum computation," Science, vol. 270, no. 5234, pp. 255–261, 1995. https://doi.org/10.1126/science.270.5234.255
[45] M.-L. Huang and T.-Y. Lin, "Dataset of breast mammography images with masses," Data in Brief, vol. 31, p. 105928, 2020. https://doi.org/10.1016/j.dib.2020.105928
[46] N. Killoran, J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook, "Strawberry fields: a software platform for photonic quantum computing," Quantum, vol. 3, p. 129, 2019. https://doi.org/10.22331/q-2019-03-11-129
[47] M. Broughton et al., "Tensorflow quantum: a software framework for quantum machine learning," arXiv preprint arXiv:2003.02989, 2020.
[48] P. Gokhale, J. M. Baker, C. Duckering, N. C. Brown, K. R. Brown, and F. T. Chong, "Asymptotic improvements to quantum circuits via qutrits," in Proceedings of the 46th International Symposium on Computer Architecture, 2019, pp. 554–566. https://doi.org/10.1145/3307650.3322253
[49] E. Farhi and N. Hartmut, "Classification with quantum neural networks on near term processors," arXiv preprint arXiv:1802.06002, 2018.
[50] S. H. Abbood, M. Rahim, and A. M. Alaidi, "DR-LL Gan: diabetic retinopathy lesions synthesis using generative adversarial network," International Journal of Online and Biomedical Engineering, vol. 18, no. 3, 2022. https://doi.org/10.3991/ijoe.v18i03.28005
[51] M. K. Abdul-Hussein, I. Obod, and I. Svyd, "Evaluation of the interference's impact of cooperative surveillance systems signals processing for healthcare," International Journal of Online and Biomedical Engineering, vol. 18, no. 3, 2022. https://doi.org/10.3991/ijoe.v18i03.28015
[52] I. A. Aljazaery, H. T. Salim, and A. H. M. Alaidi, "Encryption of color image based on DNA strand and exponential factor," International Journal of Online & Biomedical Engineering, vol. 18, no. 3, 2022. https://doi.org/10.3991/ijoe.v18i03.28021
[53] M. Al-dabag and R. Al-Nima, "Anticipating atrial fibrillation signal using efficient algorithm," International Journal of Online and Biomedical Engineering (IJOE), vol. 17, no. 2, pp. 106–120, 2021. https://doi.org/10.3991/ijoe.v17i02.19183
[54] N. Alseelawi and H. T. Hazim, "A novel method of multimodal medical image fusion based on hybrid approach of NSCT and DTCWT," International Journal of Online & Biomedical Engineering, vol. 18, no. 3, 2022. https://doi.org/10.3991/ijoe.v18i03.28011
[55] C. Wang, A theory of generalization in learning machines with neural network applications. University of Pennsylvania, 1994.

7 Authors

Amjed Y. Sahib received his B.S. and M.S. degrees in electrical engineering electronic and communication department from the University of Baghdad in 2013. He received Ph.D. degree from the college of electrical and computer engineering, communication department, university of Tabriz, Iran in 2022. His research interests artificial intelligent system (AI), finite element method (FEM), and their application in the electromagnetic system. He publishes some articles in this area.

Muazez Al Ali is the head of the department of Restorative and Prosthetic Dentistry, college of dentistry at Al Ayen University – Iraq. He obtained his master's degree in oral and maxillofacial pathology from Baghdad university (e-mail: drmuzazalali@gmail.com).

Musaddiq Al Ali is a double major in laser physics and mechanical engineering. He obtains his two master's degrees in laser physics and solid mechanics at Baghdad university. He got his Ph.D. degree from Hiroshima university on the subject of optimizing metal 3D-printed orthopedics. He works in several corporations such as Sharp and Comtech in Japan. During his work in SHARP, he designs and built the WIFI 6 compliance measurement system to evaluate the current production line of 5G based AQUAS smartphones. Currently, he is a researcher in the section of the department of advanced science and technology in smart vehicle section at Toyota technological institute (e-mail: alali@toyota-ti.ac.jp).

Article submitted 2022-12-19. Resubmitted 2023-01-17. Final acceptance 2023-01-19. Final version published as submitted by the authors.