IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 16 November 2023, accepted 7 December 2023, date of publication 14 December 2023, date of current version 20 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3343120

RESEARCH ARTICLE

Underwater Animal Identification and Classification Using a Hybrid Classical-Quantum Algorithm

SHEENA CHRISTABEL PRAVIN [1], G. ROHITH [1], V. KIRUTHIKA [1], E. MANIKANDAN [1,2], S. METHELESH [1], AND A. MANOJ [1]

[1] School of Electronics Engineering, Vellore Institute of Technology, Chennai Campus, Chennai 600127, India
[2] Centre for Innovation and Product Development, Vellore Institute of Technology, Chennai Campus, Chennai 600127, India

Corresponding author: Sheena Christabel Pravin (sheenachristabel.p@vit.ac.in)

This work wass supported by Vellore Institute of Technology, Chennai Campus, Chennai, India.

ABSTRACT Underwater Animal Identification and Classification is gaining significant importance in recent times due to the growing demand for ecological surveillance and biodiversity monitoring. Classical Deep learning techniques have been prominently used for these tasks, but due to the live capture of animals in complex environments, a limited sea-animal image dataset, and the complex topography of the seafloor, particularly in shallow waters, sediments, reefs, submarine ridges, and ship radiation, the efficacy of identification and classification is still a bottleneck for several researchers. In this paper, three hybrid Classical-Quantum neural networks ResNet50-QCNN, ResNet18-QCNN and InceptionV3-QCNN have been proposed for underwater quantum-classical Animal Identification and Classification. It significantly lessens the complexity of classical computer processing data by using quantum devices to minimize dimension and denoise datasets. The numerical simulation results demonstrate that the quantum algorithm is capable of effective dimensionality reduction and an improvement in classification accuracy. The hybrid approach offers polynomial acceleration in dimension reduction beyond classical techniques, even when quantum data is read out classically. The three hybrid models, viz., ResNet50-QCNN, ResNet18-QCNN, and InceptionV3-QCNN, displayed classification test accuracy of 88%, 80.29%, and 70%, respectively, revealing that ResNet50-QCNN performed best in identifying and classifying underwater animals.

INDEX TERMS Hybrid quantum circuit, Inceptionv3-QCNN, Resnet50-QCNN, ResNet18-QCNN, sea-animal image dataset.

I. INTRODUCTION

Seawater covers the majority of the planet's surface, although most of its volume and expected seabed are still unexplored. This is mostly a result of the unique features of the ocean environment, which make it unfriendly to people and provide practical obstacles to its exploration. These features include high pressure, low temperature, and the absence of light. However, it is now possible to dive in almost any area of the deep marine ecosystem because of advancements in robotic platforms and sensor technology [1]. Therefore, to support realistic and scientifically supported management practices, a great deal more work has to be put into acquiring baseline knowledge about marine ecosystems in terms of species and the features of their habitats [2]. The last 20 years have seen an increase in underwater imaging capabilities due to the development of High-Definition (HD) optics and the introduction of low-light technology, which, when combined with acoustic multi-beam cameras, currently allow for night vision. With the development of high-definition (HD) optics and the introduction of low-light equipment over the past 20 years, underwater imaging assets have increased and are currently capable of seeing in the dark thanks to acoustic multi-beam cameras [3]. For the extended and ongoing monitoring of marine biodiversity in any other operational scenario, the automation of image capture and processing for

The associate editor coordinating the review of this manuscript and approving it for publication was Jiajia Jiang.

animal tracking before classification is relevant [4], [5]. Even though higher-quality HD outputs can now be obtained with the use of improved imaging technologies, pre-processing is almost always required because of the significant environmental variability [6]. While the bottom is static and almost completely devoid of details in controlled laboratory environments, images are acquired in uncontrolled field settings, such as the seas, with variable environmental lighting or artificial lighting for coastal to deep-sea applications, floating particles, and variable substrates as background [7], [8], [9]. The identification and tracking of animals within the Field of Views (FOVs) and the extraction of their morphological features for classification are made more difficult by this variability [10].

Pre-processing of images is necessary to enhance the detection and posterior classification of animals. With increased interest in the field of image classification in the deep learning domain, many algorithms and methods have been introduced to effectively classify objects. Image classification has been widely applied to versatile domains. Recently, underwater image classification has given a new path to exploring the domain of computer vision. However, due to the complexity of the underwater environment, underwater photography exhibits edge and detail deterioration, poor contrast between target and background, and noise pollution [11]. Artificial Intelligence (AI) is increasingly employing Machine Learning (ML) techniques, including applications in medicine [12], [13], agriculture [14], industry [15], and marine ecology [5], [16], [17]. From the past decade, there has been a remarkable increase in the use of AI-based algorithms for the monitoring and categorization of creatures in seabed, or benthic, realms using data from cabled observatories [10], [18], [19], [20].

The categorization and identification of marine animals such as fish, plankton, and coral reefs aid in the management of marine biological systems and biodiversity, as well as the study of marine biological species differences and the conservation of endangered marine organisms. The distribution of diverse marine animals is useful for analyzing the impact of global warming and human exploitation of marine resources on marine organisms, as well as for guiding human rational use of marine resources. Several difficulties are being faced with unprocessed images from underwater [21], [22], and [23], viz., the attenuation of light that causes the images to look blurred and dull. Voluminous studies [24], [25] have been made to ratify the blurring of underwater images using various deep learning models [26], [27], [28] in different environments. Image pre-processing facilitates better training of deep learning networks. By using a classical Convolutional Neural Network (CNN), the model can be trained to achieve high classification accuracy, but it will incur high resource utilization in terms of time and space. This is where transfer learning started to gain significance.

However, when there is a large quantum of datasets, manual processing is no longer feasible. Consequently, an automated procedure is required to enhance these massive datasets, either through manual analysis or by utilizing them subsequently in a detection and classification process to yield superior outcomes. This motivated the proposal of the classical-quantum algorithm, which is founded on the theory of quantum computing that views data as qubits. A qubit is a quantum bit by definition and cannot be regarded as a parameter. A quantum bit is capable of simultaneously holding 1 and 0, as the "superposition state," and it is a fundamental aspect of the quantum world. As a result, a single qubit has the potential to participate in millions of processes at once making quantum computing extremely quick.

Quantum Computing (QC) has made major strides recently, and Quantum Machine Learning (QML) has experienced a sharp rise in popularity and productivity [29]. Due to its inherent parallelism and high speed of execution, QC would be able to ratify the errors in classification seen in traditional ML [30]. The research on quantum machine learning [31] for image categorization has revealed some positive results and shown that there is a huge potential for advancement. It is quite challenging to train a deep learning model for image classification, particularly underwater images, because of the complex environment of the ocean and the ever-present issue of image blurring. Due to the complexity of the underwater environment, underwater photography suffers from edge and detail erosion, poor target background contrast, and noise pollution. Low contrast, blur, and distortion could be more severe in underwater photography. The categorization and identification of marine organisms, including coral reefs, fish, and plankton. Plankton, fish, and coral reefs are all very varied in size.

The emphasis on hybrid models, i.e., the situation in which quantum variational circuits and conventional neural networks are concurrently trained to perform difficult computing tasks, has grown significantly. In this context, three new transfer learning variants emerge naturally, in addition to the standard Classical-to-Classical (CC) transfer learning strategy, in which some pre-acquired knowledge is transferred between classical networks: Classical to Quantum (CQ), Quantum-to-Classical (QC), and Quantum to Quantum (QQ). CQ transfer learning is especially interesting because it allows one to use any cutting-edge deep neural network to classically pre-process large input samples (such as high-resolution images) and to use a variational quantum circuit to manipulate a small number of very informative features one at a time. The strategy is highly practical since it leverages the capabilities of quantum computing in conjunction with the tried-and-true techniques of classical machine learning. However, QC and QQ transfer learning may potentially be highly intriguing strategies, particularly when big quantum computers become accessible. Here, fixed quantum circuits might be provided as generic quantum feature extractors beforehand, imitating popular classical models that are frequently employed as blocks that have already been trained like ResNet18 [49], ResNet50 [50] and inception V3 [51]. These

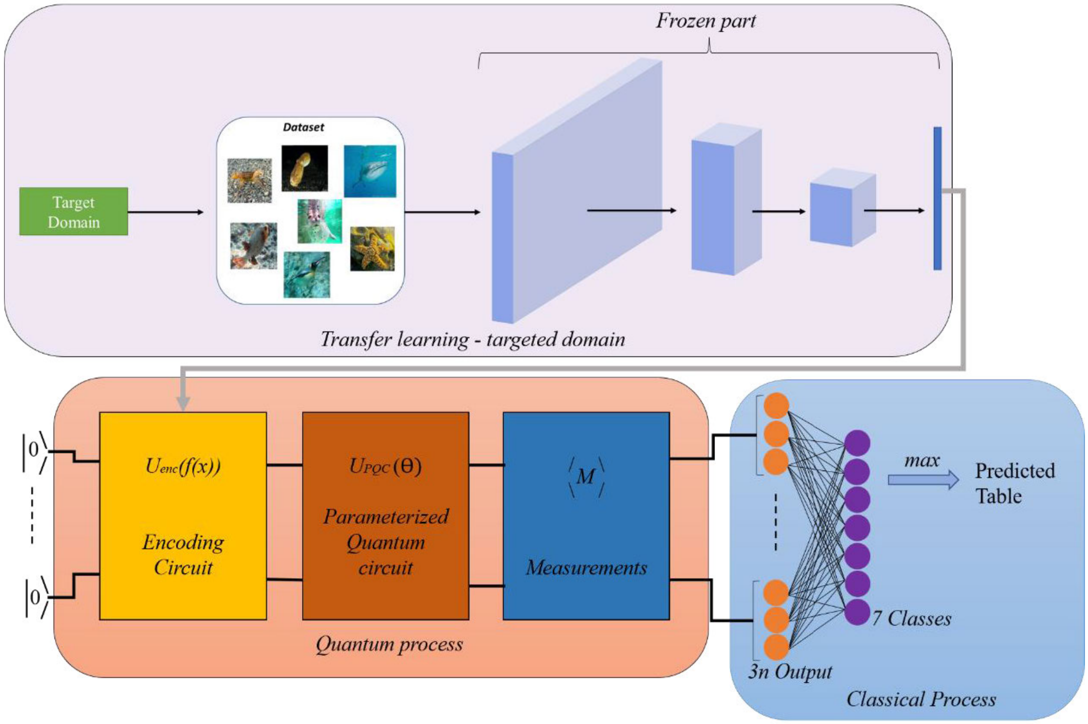


FIGURE 1. Architecture of the proposed hybrid quantum classical convolutional neural network.

traditional state-of-the-art deep networks may be applied to CC and CQ transfer learning, or they can be substituted with quantum circuits in the QC and QQ variations of the same methodology. Exactly those traditional pre-trained models are used in CQ transfer learning as feature extractors, and these features are subsequently post-processed on a quantum computer, for example, by utilizing them as input variables for the dressed quantum circuit mode.

This research work proposes the agglomeration of a quantum model with a classical model for efficient image classification, which would improve the classification process with more accuracy and reduced complexity. The authors were inspired to carry out image classification using a quantum circuit because a quantum particle, such as an electron, can exist in a state of multiple probabilities simultaneously rather than a single state with a definite position or momentum. The superposition principle may be applied to systems of multiple particles, such as atoms and molecules, which can exist in the superposition of different energy states. Studies on a wider perspective on hybrid classical-quantum models [24], [32], [33], and over versatile data sets [34].

To the best of the author's knowledge, the significant and unique contributions of this research paper are:

- Three hybrid QCNNs viz. ResNet18-QCNN, ResNet50-QCNN and Inception v3-QCNNs have been proposed for the task of multi-class underwater animal identification and classification.
- A regress comparative analysis of the classical-classical transferred models with the classical-quantum models is presented.

The rest of the paper is organized as follows-Section II concentrates on the background of literature which forms the basis for proposing the algorithm. Section III concentrates on the proposed technique. Section IV concentrates on the results and discussion including training and validation of results separately. Section V concentrates on the conclusion.

II. BACKGROUND

During the inception stage of the classification tasks, the commonly used techniques for classifying underwater images were based on image processing and pattern recognition technologies. These techniques involved preprocessing underwater images by filtering them or performing segmentation and other operations [22], [35]. A technique was presented [36] that uses support vector machines (SVM) to train fish images to lessen the effects of different disturbances in the underwater environment. The technique [36] provided an increased classification accuracy of 74.8% in the data set of 15 fish species with a total of 24,000 images. This gave rise to Artificial intelligence [37] and convolutional neural networks (CNNs) [38] and many new and effective techniques for classifying images have been added The benefit of employing CNN for image classification is that image features can be extracted and filtered automatically, saving time and effort. This can be automatically finished by the convolution operation. Neural networks can generate higher semantic-level features for classification as convolution is deepened. To identify fish in coral reefs, Villon proposed a deep-learning classification technique based on CNN [39].

To realize the function of fish recognition and detection, Labao et al. [40] developed a set of fish recognition and detection systems combined with a long short-term memory network and a convolutional neural network based on region and tested it on eighteen field-captured video data. Guo et al. [41] completed the identification of sea cucumber with the highest accuracy of 89.53% by using the depth residual network. The residual network was incorporated into the CNN network by Prasetyo et al. [42], who also suggested a VGGNet with a multi-level residual MLR-VGGNet. It integrates the deep, advanced features while keeping the primary and intermediate features from the early convolution blocks. MLRVGGNet's classification accuracy on the FishGres and Fish4Knowledge datasets is 99.69%. Gómez-Ríos et al. [43] constructed a two-level classifier using three CNN models to classify the structure, shape, and texture of coral, in addition to introducing the residual network. Ananda et al. [44] used ResNet152 to classify and detect brain images after transfer learning. Furthermore, an attention mechanism is introduced into the network to train it to assign weights to different features, pay attention to more important features, and ignore secondary features [45]. Alshdaifat et al. [46] employed the example segmentation approach to achieve 95.2% accuracy on the Fish4Knowledge dataset after adjusting the brightness of underwater fish footage to reduce blur. Using many sets of coral data, Ganesan and Santhanam [47] optimized the random parameters produced by ELM using the chimpanzee optimization technique, achieving 95% to 98% classification accuracy. The method still has an issue with local optimization, and manually changing super parameters results in feature duplication that cannot be avoided. Extracting and classifying picture characteristics for the visual categorization of marine species is difficult because of the intricacy of the living environments of marine organisms and the limited resolution of underwater imaging technologies. From the literature survey in [48] and literatures [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], and [47], the following are the research gaps identified and gratified as the first of its kind using the proposed technique.

- Hybridization of both a classical model as well as a quantum model imbibing the notion of transfer learning for underwater animal identification and classification.
- The proposed algorithm facilitates to identify significant characteristics and extract useful elements from noisy underwater and different illumination conditions of the pictures.
- This model is suitable to handle large amount of the datasets as the quantum computing is used.

III. PROPOSED TECHNIQUE

In this research work, the Classical-Quantum transfer learning approach is adopted. This research work has two phases, the first phase is the classical Convolution Neural Network (CNN) network performs feature extraction followed by the second phase is the classification by the Quantum circuit. In the first phase, for the classical part, the ResNet18, ResNet50 and Inception V3 were used without any changes introduced to the architecture of their skeleton models. Varied classification accuracies and computational complexities resulted in making the choice of the best classical generative model. In the second phase, for the Quantum part, Hybrid QCNN is appended where the Parameterized Quantum Circuit was appended to the model alongside after the detachment of the final layers of the pre-trained classical models for the categorical classification of the images into 7 different classes. The underlying assumption of this approach is that a classical model can still be an efficient feature extractor for other problems even if it has been optimized for one particular task. In the context of transfer learning, when some of the last layers were eliminated, model A which was trained for a specific purpose would work as a feature extractor. The latter layers of a network are frequently tailored to the particular task, whereas intermediate characteristics were more generic and better suited for transfer learning, which is why the final layers of the classical model were discarded. Fig.1 illustrates the workflow of the proposed model.

A. CLASSICAL CONVOLUTION NEURAL NETWORK

Three classical networks have been introduced to take the role of a generative model viz. ResNet18 [49], ResNet50 [50] and inception V3 [51].

ResNet18 [49] is a specific type of CNN architecture and a smaller version of the original ResNet architecture, with 18 layers, making it computationally efficient. The key innovation of the ResNet architecture is the use of a building block called a residual block, where the input is first passed through one or more convolutional layers, and the output of these layers is then added to the original input, called a *shortcut* connection. This allows the network to learn the residual mapping from the input to the output, rather than the full mapping, making it possible to train deeper networks sans the problem of vanishing gradients.

ResNet50 [50] is a variant of the CNN architecture for deep learning, which belongs to the ResNet family of architectures. Residual connections are skip connections that bypass one or more layers in the network, allowing gradients to flow more easily during training and helping to alleviate the vanishing gradient problem. This enables the training of deeper networks with hundreds or even thousands of layers, which can learn more complex representations from input data. ResNet-50 consists of 50 convolutional layers, with additional batch normalization, ReLU activation, and pooling layers. The residual connections are introduced in pairs of two or more layers, enabling the network to learn residual mappings. The architecture also includes a pooling layer followed by a fully connected layer at the end for classification or other tasks.

Inception V3 [51] is a variant of the CNN architecture for deep learning which is built of inception modules with multiple convolutional layers with different filter sizes

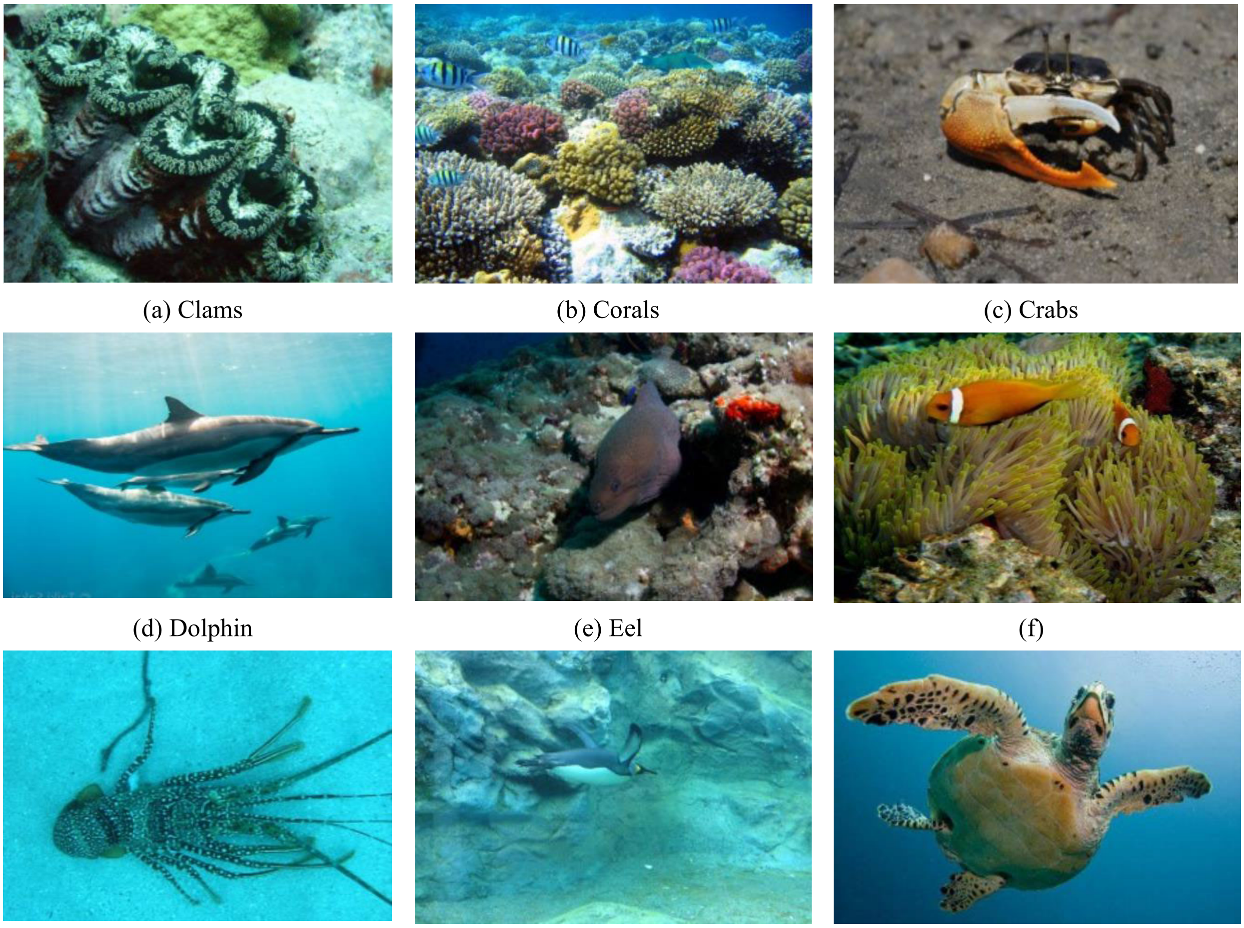

(a) Clams (b) Corals (c) Crabs (d) Dolphin (e) Eel (f) (g) Lobster (h) Penguin (i) Turtle Tortoise

FIGURE 2. Sample images from the sea animal dataset.

(1×1, 3×3, and 5×5), as well as max pooling and average pooling layers. They are constructed to capture information at different scales and resolutions in parallel. These inception modules consist of multiple convolutional layers.

The outputs of these different operations are then concatenated together to form the final output of the inception module. This allows the network to capture both local and global contextual information, making it more effective at learning complex features in images.

In this experiment, the ResNet18 classical pre-trained model was introduced, discarding its final layers; which makes the ResNet18 model a generative network. A classical layer was then introduced to classify the images into different classes based on the extracted features. Further, other CNN variants viz. ResNet80, ResNet50 and Inception V3 were tested for their candidature to be a suitable generative model. In this research work, pure classical models viz. ResNet18, ResNet50 and Inception V3 were trained on *Sea Animal*image dataset. Also, the proposed hybrid quantum models viz. ResNet50-QCNN, ResNet18-QCNN and Inception V3-QCNN were trained over the same dataset, implemented and tested for their accuracy.

B. CLASSICAL-QUANTUM TRANSFER LEARNING

The practice of translating information from a classical machine-learning model to a quantum machine-learning model is known as *Classical-Quantum transfer learning*. The objective is to enhance the training and performance of a quantum machine learning model by using the learned classical characteristics, representations, or models. The idea behind *Classical-Quantum transfer learning*is that classical machine learning models can sometimes learn high-level features or representations that are relevant to a quantum machine learning task. By using these learned features or representations as input to a quantum machine learning model, there is potential scope for improving the efficiency and accuracy of the quantum model.

Using the pre-trained classical model as a feature extractor and the extracted features as input to a quantum machine learning model is one method of performing Classical-Quantum transfer learning. Consider the scenario where a traditional machine learning model has been trained to categorize pictures of handwritten digits. This model can be used to extract sophisticated information from the images, such as texture analysis, edge identification, and shape recognition. The quantum machine learning model that has been trained to categorize handwritten digits can then be fed with these extracted attributes. Inspired by the literature in [52], the trained classical models were used as feature extractors by removing the end layers of the models and then combining them with a quantum circuit which is the Parameterized Quantum Circuit. This combination results in the Dressed Quantum Circuit. A Quantum circuit-based classification may be understood as an example of classification between Class A and B which is represented by values 1 and 0 respectively.

1) PARAMETERIZED QUANTUM CIRCUIT

One of the most fundamental quantum algorithms is a variational or parameterized quantum circuit. The variational quantum circuit is a hybrid technique that takes advantage of the pros of both quantum and classical computations. It is a kind of quantum circuit with configurable parameters that are iteratively tuned by a classical computer. These parameters can be thought of as artificial neural network weights.

A quantum layer may be defined as a unitary operation when the input state is altered or changed x of nq quantum subsystems which are like qubits, which return an output state as $|y\rangle$ is obtained.

$$\mathrm{L} \,:\, |\mathrm{x}\rangle \rightarrow |\mathrm{y}\rangle = \mathrm{U}(\mathrm{w})|\mathrm{x}\rangle \tag{1}$$

Here in equation (1), w is the array of classical variational parameters. A variational quantum circuit with depth q is the result of the product of many unitaries with different weights, or many quantum layers:

$$\mathrm{Q} = \mathrm{Lq} \circ \bullet \bullet \bullet \circ \mathrm{L2} \circ \ \mathrm{L1} \tag{2}$$

A real vector $\mathbf{x}$ is imbibed into a quantum state $|x\rangle$ to introduce classical data into a quantum network. A variational layer built on the input vector is applied to a reference state.

$$\mathrm{E} \,:\, x \rightarrow |x\rangle = E(x)|0\rangle \tag{3}$$

Instances include rotations with a qubit or single mode displacements parameterized. The embedding layer maps a classical vector to a quantum Hilbert, in contrast to the matrix layer.

By measuring the expected value of n local observable, the quantum circuit may be used to extract the classical data $\mathbf{y}$. $\hat{y} = [\hat{y}_1, y_2, \ldots, y_{nq}]$. This procedure, which converts the quantum state into a classical vector, is known as a measurement layer:

$$\mathrm{M} : |x\rangle \rightarrow y = \langle x|\hat{y}|x\rangle \tag{4}$$

For example, there are two states 1 and 0 referring to two classes, this measuring expectation value is done by checking which side the value tends to be close to; either 0 or 1 (the more towards one state, the better the result) and then choose the corresponding class. This can also be increased to more than two classes.

The complete quantum network, which includes the preliminary embedding layer and the end measurement layer, can be expressed as follows:

$$Q = \mathrm{M} \circ \mathrm{Q} \circ \mathrm{E} \tag{5}$$

As quantum computing will be buried between the quantum circuit layers, if seen from a broad perspective, Q is just a black box compared to the classical deep learning network. If viewed from a shallow perspective, this will appear to map from one classical feature vector to another.

2) DRESSED QUANTUM CIRCUIT

One of the major contributions of this research work is the application of transfer learning to the *Classical-Quantum* interface. To accomplish this, a classical neural network must be connected to the quantum neural network. Thus the resulting model is referred to as the *Dressed Quantum Circuit*. The basic structure of a *Dressed Quantum Circuitis* the quantum layer, sandwiched between two classical layers. To a variational circuit and its $\mathrm{n_q}$ subsystems, signal conditioning of the input and output data was introduced by adding a classical layer at the beginning of the quantum network, creating the *Dressed Quantum Circuit*:

$$\hat{\mathrm{Q}} = \mathrm{L}_{\mathrm{nq}\rightarrow\mathrm{nout}} \circ \mathrm{Q} \circ \mathrm{L}_{\mathrm{nin}\rightarrow\mathrm{nq}} \tag{6}$$

where $\mathrm{L}_{\mathrm{nq}\rightarrow\mathrm{nout}}$ indicates the number of inputs and Q is the bare quantum circuit defined in Eq. (6). The major computing is carried out by the quantum circuit Q in this example, as opposed to a complicated hybrid network where collaborating classical and quantum processors divide the workload. The classical layers are primarily in charge of data embedding and readout.

Rewriting the above equation based on the input

$$\bar{\mathrm{Q}} = \mathrm{L}_{4\rightarrow 2} \circ \mathrm{Q} \circ \mathrm{L}_{2\rightarrow 4} \tag{7}$$

The selected embedding map sets up each qubit in a balanced superposition of |0> and |1> and then rotates it around the Bloch sphere's y-axis using a classical vector x as its parameter:

$$\varepsilon\,(x) = \oplus(R_y\left(\frac{x_k\pi}{2}\right)H)|0000\rangle \tag{8}$$

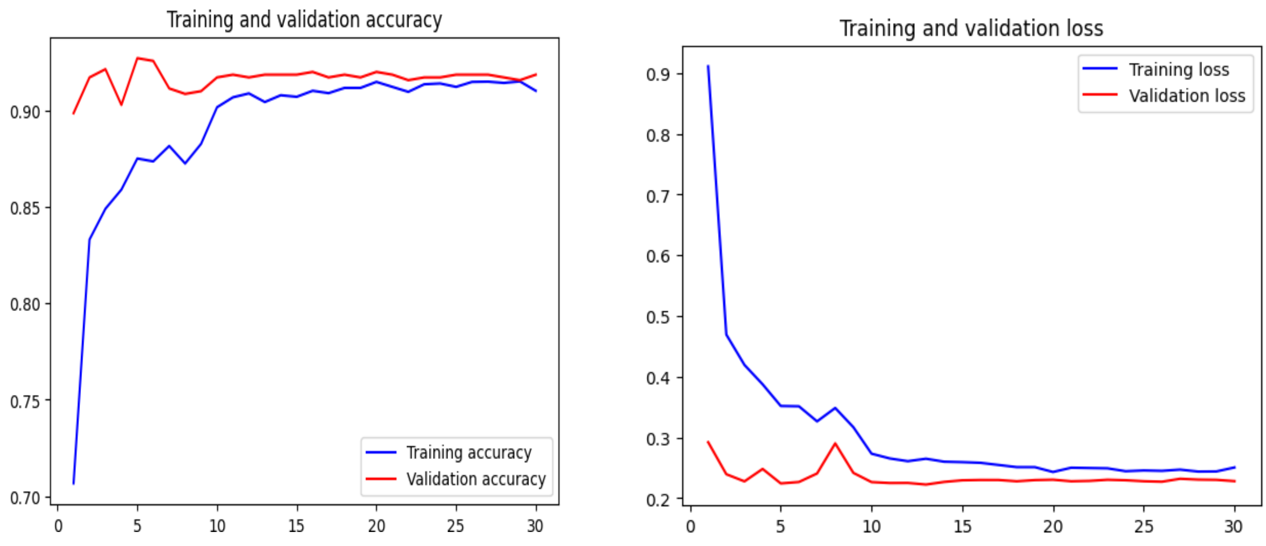
where H is called the single-qubit Hadamard gate. There are around 5 variational layers in the trainable circuit which are represented by

$$L\,(w) : |x\rangle \rightarrow |y\rangle = K \oplus (w_k)\,|x\rangle \tag{9}$$

k is a unitary entangling operation composed of three controlled-NOT gates.

Each classical data was multiplied and rotated. To finally classify the data, a measurement layer was added at the end (i.e.) an expectation value of the $Z = diag(-1, 1)$.

This produces a classical output vector, suitable for additional post-processing. The classification is carried out using argmax(y), where $y = (y_1, y_2)$ is the output of the *Dressed Quantum Circuit*, given an input point with coordinates $x = (x_1, x_2)$.



(a) (b)

FIGURE 3. **An output of ResNet18 model's binary classification training and validation accuracy and loss using classical transfer learning for 30 epochs indicating drastic variations.**

(a) (b)

FIGURE 4. **An Output of ResNet50 model's binary classification training and validation accuracy and loss using classical transfer learning for 30 epochs indicating drastic variations.**

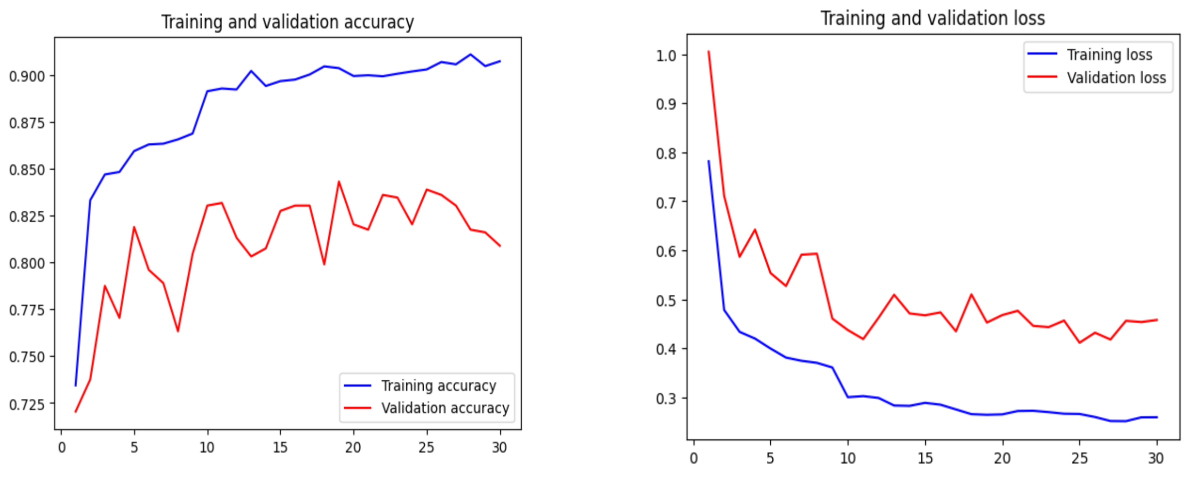
IV. RESULTS AND DISCUSSION

The results and discussion are presented in three sections- A. Dataset Description, B. Training and Testing of the results for both classical and classical Quantum, and C. Validation of the results for both classical and classical Quantum.

A. DATASET DESCRIPTION

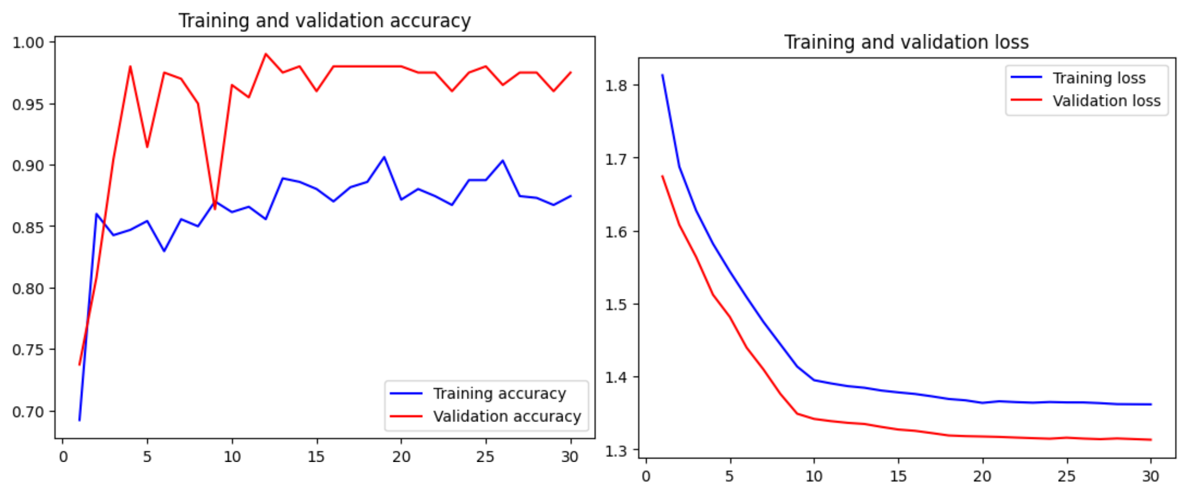
The majority of living forms evolved in watery environments. In terms of volume, the seas supply around 90% of the world's living area. Fish, which can only be found in water, were the first vertebrates discovered. Some of them evolved into amphibians, which spend portions of the day on land and in water. A few amphibian subgroups, which included sea turtles, seals, manatees, and whales, evolved into reptiles and mammals. Plant life that grows in the water, such as kelp and other algae, supports certain underwater ecosystems. Phytoplankton, which are key primary producers, provide the foundation of the ocean food chain. The dataset used for this can be found on the following website- https://www.kaggle.com/datasets/
vencerlanz09/sea-animals-image-dataste

The collection includes several images of aquatic creatures. Some images were obtained from pixabay.com and do not require a license or attribution to be used. There are now 19 distinct classes offered, which may be expanded in the future. The 19 distinct classes include Seahorse,



(a) (b)

FIGURE 5. **An output of ResNet18 model's binary classification training and validation accuracy and loss using classical quantum transfer learning for 30 epochs indicating drastic variations.**



(a) (b)

FIGURE 6. **An output of ResNet18 model's binary classification training and validation accuracy and loss using classical quantum transfer learning for 30 epochs indicating drastic variations.**

Nudibranchs, Sea Urchins, Octopus, Puffers, Rays, Whales, Eels, Crabs, Squid, Corals, Dolphins, Seals, Penguin, Starfish, Lobster, Jelly Fish, Sea Otter, Fish, Shrimp and Clams. Fig.2 shows the sample of images representing the dataset. The images were adjusted to *(300px, n)* or *(n, 300px)*, where *n* is less than 300px.

The proposed Hybrid QCNNs viz. Resnet18-QCNN, Resnet50-QCNN, and the Inception v3-QCNN models were trained over the open-source dataset containing 7 classes of underwater organisms with each class containing nearly 500 images. The seven distinct classes utilized were Crabs, Dolphins, Fish, jellyfish, Lobster, Sea Urchins, and starfish. It is analyzed from the dataset that 7 distinct classes chosen from 19 is because

(i) The balanced dataset and the clarity of the image were evident only in the 7 distinct classes.

(ii) Moreover, among the 19 classes in the dataset, 7 distinct classes could be grouped by covering the rest 12 and hence utilized for the proposed technique.

After data augmentation, the total images were populated to 7500. When the images were given as input to the classical network (i.e.) the feature extractor could discern unique features of the underwater images. The extracted images were given as input to the PQC.

B. TRAINING AND TESTING OF THE RESULTS

1) CLASSICAL-CLASSICAL TRANSFER LEARNING USING CONVOLUTION NEURAL NETWORK

Three classical networks have been introduced to take the role of a generative model viz. ResNet18, ResNet50 and inception V3. First, the ResNet18 classical pre-trained model was introduced, discarding its final layers; which makes the

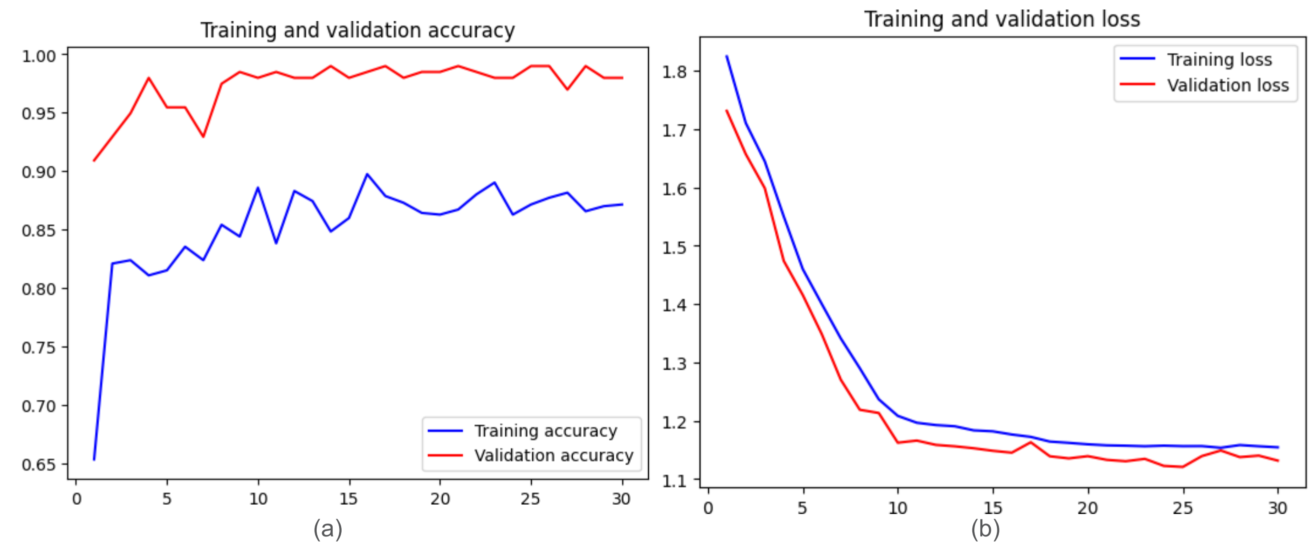


FIGURE 7. **An output of ResNet50 model's binary classification training and validation accuracy and loss using classical quantum transfer learning for 30 epochs indicating drastic variations.**

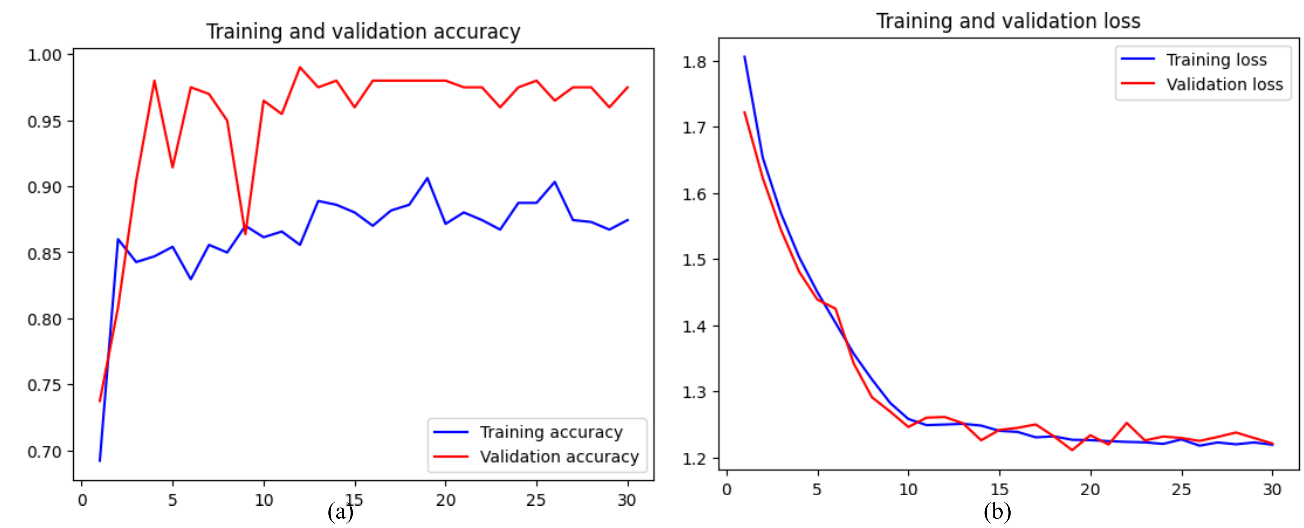


FIGURE 8. **An output of ResNet18 model's binary classification training and validation accuracy and loss using classical quantum transfer learning for 30 epochs indicating drastic variations.**

ResNet18 model a generative network. A classical layer was then introduced to classify the images into different classes based on the extracted features. It was observed that the model could classify well despite being trained over a sparse dataset.

Initially, the dataset was tested for binary classification using pure Classical-Classical transfer learning, the ResNet18 excelled in performance by displaying a test accuracy of 98% with the following hyper-parameters: step size: 0.0004, batch size: 4 and the number of epochs of 100. The ResNet50 model exhibited a test accuracy of 99% but at the cost of an additional 20 minutes of run time to complete the training process. The Inception v3 deep learning classical model required an additional 132 minutes than the ResNet18 model and rendered a test accuracy of 90%. Thus, for binary classification, the ResNet18 model was observed to yield the maximum classification accuracy at reduced computational complexity. On the other hand, binary classification has also been experimented with using Classical-Quantum hybrid models. The proposed ResNet18-QCNN could render a test accuracy of 99% with an increase in the number of epochs to 30. Upon successful results, the following seven distinct classes utilized: Crabs, Dolphins, Fish, Jelly Fish, Lobster, Sea Urchins, and starfish were used to train the ResNet18, ResNet 50 and Inception V3-hybrid QCNN models.Fig.3 to Fig.5 shows the training and its corresponding loss that occurred due to the classical transfer learning approach. Data augmentation in the form of Horizontal flip, 25% grayscale, Hue between 100° and +100°, saturation between −70% and +70%, brightness between −40% and +40%, exposure between −25% and +25%, blur up to 2.25px and noise up to 15% of pixels was adopted to increase the image data. The Classical binary classification model for the two classes mentioned above was trained with a learning rate of 0.004, a batch size of 4, with a reduction in learning rate of a factor of 0.1 every 10 epochs for 100 epochs. In the second phase of this research work, 7-class classification was undertaken with the transfer-learned ResNet18, ResNet50, and the inception v3 models. 7 classes in the same open-source dataset having a total of around 8422 images and trainable images being 7371 images. The same training procedure was undertaken to transfer-learn the above-cited three classical models to yield a 7-class classification.

The hyper-parameters used for this phase of classification include a batch size of 256, with a learning rate of about 0.001 and 30 being the number of epochs. The ResNet18 model displayed a classification accuracy of 91.71% over the 7-class classification, which can be observed from Fig.3 with the computation complexity of 386 minutes and 32 seconds. ResNet50 model was then trained similar to the ResNet 18 model. It holds the same architecture as that of the ResNet18 with just an increased number of layers. The same dataset with the retained hyper-parameters was used with ResNet50 as well. The computation complexity of the model is 926 minutes and 12 seconds resulting in a classification accuracy of 94% which is by far the highest accuracy, which can be observed from Fig.4. A similar training procedure was considered for the inception V3 transfer learning classical model. The same procedure was then repeated with the Inception v3 model for the same hyper-parameters which gave a test accuracy of 81.97%, which can be observed from Fig.5. The computation complexity of the model is 1623 minutes and 14 seconds.

2) CLASSICAL-QUANTUM TRANSFER LEARNING USING CONVOLUTION NEURAL NETWORK

As in the classical transfer learning, the seven distinct classes utilized were Crabs, Dolphins, Fish, Jellyfish, Lobster, Sea Urchins, and starfish were used to train the Resnet18, Resnet 50 and Inception V3 models. Fig.6 to Fig.8 show the training and its corresponding loss that occurred due to the classical-quantum transfer learning approach (hybrid QCNN models). The hyper-parameters were chosen as follows on regress experimentation: a learning rate of 0.001 with a batch size of 256 and the number of epochs as 100. These hyper-parameters enabled the model to give the highest accuracy with the hyper-parameters mentioned above. The model's training was completed in a total duration of about 515 minutes and 56 seconds and got the results for test accuracy of 80.29%, which can be seen in Fig.6. The proposed ResNet50-QCNN Classical-Quantum transfer learning model was trained to retain the hyperparameters used for ResNet18-QCNN model. It was observed that the ResNet50-QCNN model took longer time for about a total duration of 1088 minutes and 44 seconds to complete the training process and rendered a test accuracy of 86.88%, which can be seen from Fig.7. The hyper-parameters set for ResNet18-QCNN and ResNet50-QCNN was adopted for the third proposed hybrid model the Inception v3-QCNN. The model was observed to run for about 1443 minutes and 5 seconds to get trained. This was observed to be the longest duration a model took in this research study. Further, the test classification accuracy dipped to 70%, which can be seen from Fig.8.

Compared to the Classical-Classical transfer learned model, the Classical-Quantum models performed better for multi-class classification in terms of increased classification accuracy and stability, as shown in both the validation loss and accuracy graphs in Fig.6 to Fig.8.

C. VALIDATION OF THE RESULTS

1) CLASSICAL-CLASSICAL TRANSFER LEARNING USING CONVOLUTION NEURAL NETWORK

Among the experimented classical models, ResNet50 was observed to perform the best with a test accuracy of 94% than the ResNet18 model which showed 91.71% and the Inception v3 which exhibited 81.71%.

However, it was observed that the ResNet18 model rendered lesser run time compared to both of the other classical deep learning models for multi-class classification. Besides, classical-quantum models viz. ResNet18-QCNN, ResNet50-QCNN, and Inception v3-QCNN were also experimented over 7-class classification. In our observation, ResNet50-QCNN displayed the highest test classification accuracy of 88% followed by ResNet18-QCNN with an accuracy of 80.29% and the Inception v3-QCNN with an accuracy of 70%. However, as anticipated, the ResNet18-QCNN model rendered reduced computational complexity than the other two proposed hybrid classical-quantum models.

Table. 1 lists the hyper-parameters set for training the pure classical models for the task of binary classification and their test accuracy thereof. The proposed models were evaluated by plotting the confusion matrix heat maps as well for multi-class classification. The confusion matrix plot for ResNet50 pure classical model was observed to show larger number of classes diagonally, which proved that the ResNet50 model performed the best for multi-class classification. Similarly, the ResNet50-QCNN had the least of the images classified erroneously as shown in the heatmap in Fig.9.

Table 2 lists the hyper-parameters set to train the Classical-Quantum transfer learning binary classification model. A new hyper-parameter namely the quantum depth was additionally introduced, which refers to the number of qubits or quantum bits involved in the quantum computation or quantum algorithm. In quantum computing, qubits are the basic units of quantum information, and the number of qubits used in the computation is indeed a critical factor that determines the computational power and complexity of the quantum algorithm.

Quantum depth was set to 6 for all the proposed C-Q models, while the rest of the hyper-parameters were retained the same set for other models. It was observed that in the hybrid quantum model, the test accuracy of

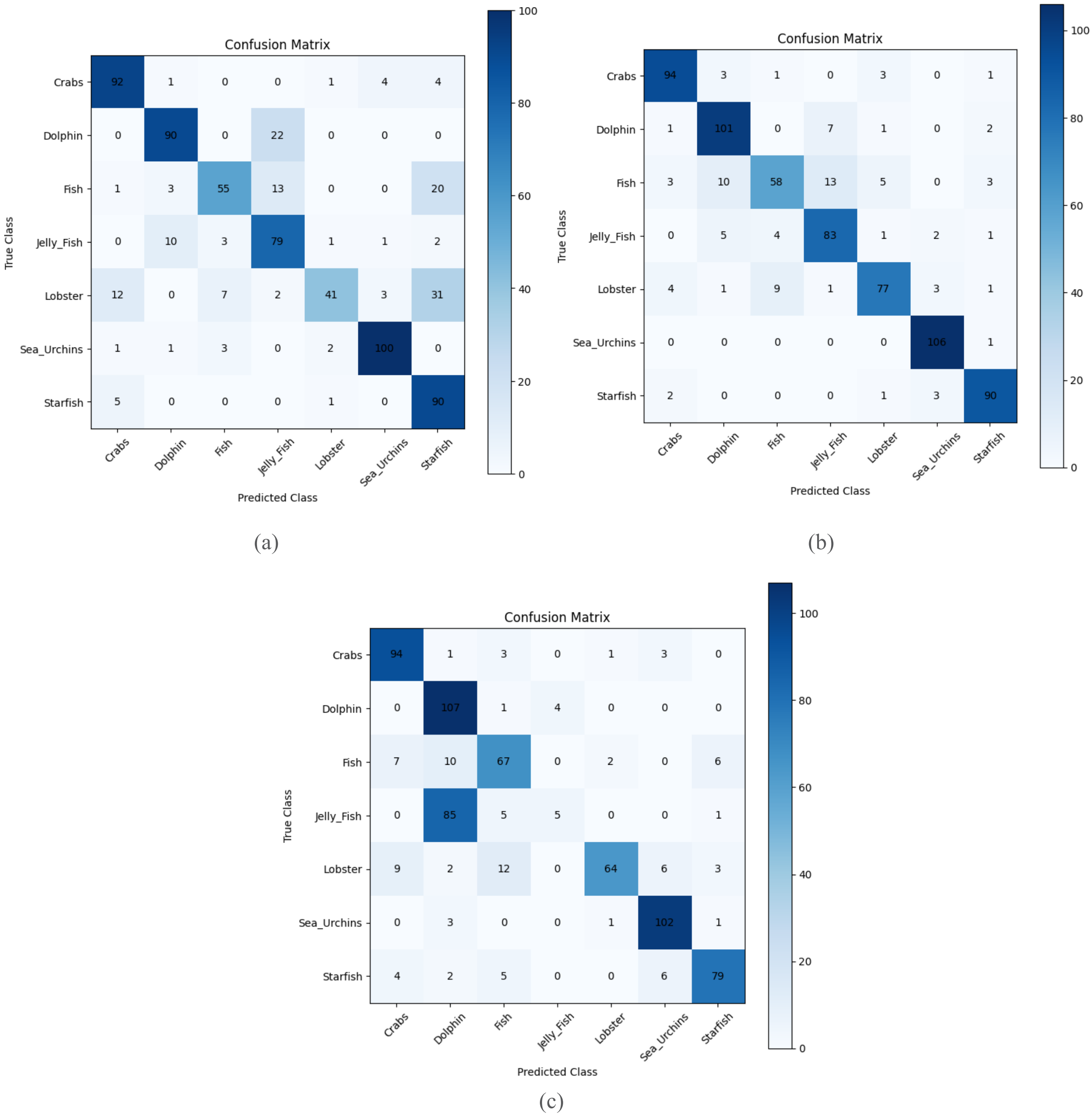


FIGURE 9. A graphical representation of the performance of all 3 transfer learning models for classical-quantum model.

ResNet18-QCNN and ResNet50-QCNN were almost similar and higher when compared to their classical counterpart. The Inception V3- QCNN model did not perform well both in terms of accuracy and computational efficiency.

Based on the above interpretations, it is concluded that

- The proposed ResNet50-QCNN and ResNet50 pure classical model performed the best in multi-class classification as against the ResNet18 pure classical model and ResNet18-QCNN which performed the best in binary classification.
- Comprehensively, the authors could decipher that the classical-quantum models outperformed the classical-classical networks in terms of test accuracy for binary classification.
- However, classical-classical models performed subtly higher than the classical-quantum models in multi-classification with respect to test classification accuracy.
- Further, in terms of computational complexity, the classical models rendered reduced computational complexity than the classical-quantum models.

TABLE 1. Hyper-parameters of all three classical-classical multi-class classification models.

CC multi-class classification model			
Hyper-parameters	ResNet 18	ResNet 50	**Inception V3**
Training batch	7371	7371	**7371**
Batch size	256	256	**256**
Duration	386m 32s	926m 12s	**1623m 14s**
Accuracy	**91.71%**	**94%**	**81.71%**

TABLE 2. Hyper-parameters of all three transfer learning model for classical-classical quantum classification model.

Hyper-parameters	**ResNet18**	**ResNet50**	**Inception V3**
Classical depth	1	1	**1**
Quantum Depth	8	8	**8**
Training batch	7371	7371	**7371**
Batch size	256	256	**256**
Duration	515m 56s	1088m 44s	**1443m 5s**
Accuracy	**80.29%**	**88%**	**70%**

- The classical-quantum models comfortably performed well over binary classification owing to lesser complexity.
- Nonetheless, the C-Q models performed marginally well in multi-class classification. This may be related to the superposition theory of quantum computation, wherein a quantum particle is expected to be in two or more states at the same time until it is measured or observed. Besides, the generalization is found to more prominent in classical-quantum models.

There are several challenges associated with Classical- Quantum transfer learning, includes

- Need to map the classical features or representations to a quantum form
- Difficulty of implementing the classical model on a quantum computer, and the need to ensure that the transferred knowledge is relevant and useful for the target quantum machine learning task.

Despite these difficulties, the proposed classical-quantum transfer learning approach has the promise of enhancing the performance of quantum machine learning models, especially in cases where classical machine learning models have already attained the best results possible on a comparable task. This is observed from the training of the model that the proposed model could classify well despite being trained over a sparse dataset.

V. CONCLUSION

Hybrid QCNNs viz. ReNet18-QCNN, ResNet50-QCNN and Inception v3-QCNN were proposed for the task of binary and multi-class image classification. The publicly available Sea Animal dataset was used to train the proposed models after data augmentation. A regress comparative experimental study over pure classical and hybrid classical-quantum models the tasks of binary and multi-class classification. The outcomes of the numerical simulation show that the quantum method may effectively reduce dimensionality and increase classification accuracy. Even when quantum data is read out traditionally, the hybrid method provides polynomial acceleration in dimension reduction beyond classical procedures. When compared to their classical counterparts, the test accuracy of ResNet18-QCNN and ResNet50-QCNN in the hybrid quantum model was shown to be both greater and almost the same. The accuracy and computational efficiency of the Inception V3-QCNN model were not good. Even when quantum data is read out traditionally, the hybrid method provides polynomial acceleration in dimension reduction beyond classical procedures. The three hybrid models, ResNet50-QCNN, ResNet18-QCNN, and InceptionV3-QCNN, performed best in recognizing and categorizing underwater species, with classification test accuracy of 88%, 80.29%, and 70%, respectively. In future, the authors will endeavor to perform quantum error correction which will eventually reduce the subtle classification inaccuracies observed in the proposed hybrid QCNNs.

REFERENCES

[1] V. Lopez-Vazquez, J. M. Lopez-Guede, D. Chatzievangelou, and J. Aguzzi, "Deep learning based deep-sea automatic image enhancement and animal species classification," *J. Big Data*, vol. 10, no. 1, pp. 1–26, Mar. 2023.

[2] R. Danovaro, E. Fanelli, J. Aguzzi, D. Billett, L. Carugati, C. Corinaldesi, A. Dell'Anno, K. Gjerde, A. J. Jamieson, S. Kark, and C. McClain, "Ecological variables for developing a global deep-ocean monitoring and conservation strategy," *Nature Ecology Evol.*, vol. 4, no. 2, pp. 181–192, Feb. 2020.

[3] J. Aguzzi, D. Chatzievangelou, S. Marini, E. Fanelli, R. Danovaro, S. Flogel, N. Lebris, F. Juanes, F. C. De Leo, J. Del Rio, and L. Thomsen, "New high-tech flexible networks for the monitoring of deep-sea ecosystems," *Environ. Sci. Technol.*, vol. 53, no. 12, pp. 6616–6631, Jun. 2019.

[4] F. Kratzert and H. Mader, "Fish species classification in underwater video monitoring using convolutional neural networks," Inst. Water Manag., Hydrol. Hydraulic Eng., Univ. Natural Resour. Life Sci., Vienna, Austria, Tech. Rep., 2018.

[5] T. Rimavicius and A. Gelzinis, "A comparison of the deep learning methods for solving seafloor image classification task," in *Proc. Int. Conf. Inf. Softw. Technol.*, 2017, pp. 442–453.

[6] T. Schoening, A. Purser, D. Langenkämper, I. Suck, J. Taylor, D. Cuvelier, L. Lins, E. Simon-Lledó, Y. Marcon, D. O. B. Jones, T. Nattkemper, K. Köser, M. Zurowietz, J. Greinert, and J. Gomes-Pereira, "Megafauna community assessment of polymetallic-nodule fields with cameras: Platform and methodology comparison," *Biogeosciences*, vol. 17, no. 12, pp. 3115–3133, Jun. 2020.

[7] J. A. Garcia, V. Sbragaglia, D. Masip, and J. Aguzzi, "Long-term video tracking of cohoused aquatic animals: A case study of the daily locomotor activity of the Norway Lobster (*Nephrops norvegicus*)," *J. Vis. Exp.*, no. 146, Apr. 2019, Art. no. e58515.

[8] M. Sokolova, A. Mompó Alepuz, F. Thompson, P. Mariani, R. Galeazzi, and L. A. Krag, "A deep learning approach to assist sustainability of demersal trawling operations," *Sustainability*, vol. 13, no. 22, p. 12362, Nov. 2021.

[9] M. Sokolova, F. Thompson, P. Mariani, and L. A. Krag, "Towards sustainable demersal fisheries: NepCon image acquisition system for automatic nephrops norvegicus detection," *PLoS ONE*, vol. 16, no. 6, Jun. 2021, Art. no. e0252824.

[10] J. Osterloff, I. Nilssen, J. Järnegren, T. Van Engeland, P. Buhl-Mortensen, and T. W. Nattkemper, "Computer vision enables short- and long-term analysis of *Lophelia pertusa* polyp behaviour and colour from an underwater observatory," *Sci. Rep.*, vol. 9, no. 1, pp. 1–12, Apr. 2019.

[11] J. Yang, M. Cai, X. Yang, and Z. Zhou, "Underwater image classification algorithm based on convolutional neural network and optimized extreme learning machine," *J. Mar. Sci. Eng.*, vol. 10, no. 12, p. 1841, Dec. 2022.

[12] J. C. Bjerring and J. Busch, "Artificial intelligence and patient-centered decision-making," *Philosophy Technol.*, vol. 34, no. 2, pp. 349–371, Jun. 2021.

[13] J.-P. O. Li, H. Liu, D. S. Ting, S. Jeon, R. P. Chan, J. E. Kim, D. A. Sim, P. B. Thomas, H. Lin, Y. Chen, and T. Sakomoto, "Digital technology, telemedicine and artificial intelligence in ophthalmology: A global perspective," *Prog. Retin. Eye Res.*, vol. 82, May 2021, Art. no. 100900.

[14] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Comput. Electron. Agricult.*, vol. 151, pp. 61–69, Aug. 2018.

[15] J. Lee, H. Davari, J. Singh, and V. Pandhare, "Industrial artificial intelligence for industry 4.0-based manufacturing systems," *Manuf. Lett.*, vol. 18, pp. 20–23, Oct. 2018.

[16] J. Aguzzi, C. Costa, K. Robert, M. Matabos, F. Antonucci, S. K. Juniper, and P. Menesatti, "Automated image analysis for the detection of benthic crustaceans and bacterial mat coverage using the Venus undersea cabled network," *Sensors*, vol. 11, no. 11, pp. 10534–10556, Nov. 2011.

[17] A. Mahmood, M. Bennamoun, S. An, F. Sohel, F. Boussaid, R. Hovey, G. Kendrick, and R. B. Fisher, "Automatic annotation of coral reefs using deep learning," in *Proc. OCEANS MTS/IEEE Monterey*, Sep. 2016, pp. 1–5.

[18] S. Marini, E. Fanelli, V. Sbragaglia, E. Azzurro, J. Del Rio Fernandez, and J. Aguzzi, "Tracking fish abundance by underwater image recognition," *Sci. Rep.*, vol. 8, no. 1, p. 13748, Sep. 2018.

[19] M. Martin-Abadal, A. Ruiz-Frau, H. Hinz, and Y. Gonzalez-Cid, "Jellytoring: Real-time jellyfish monitoring based on deep learning object detection," *Sensors*, vol. 20, no. 6, p. 1708, Mar. 2020.

[20] A. Zuazo, J. Grinyó, V. López-Vázquez, E. Rodríguez, C. Costa, L. Ortenzi, S. Flögel, J. Valencia, S. Marini, G. Zhang, H. Wehde, and J. Aguzzi, "An automated pipeline for image processing and data treatment to track activity rhythms of *Paragorgia arborea* in relation to hydrographic conditions," *Sensors*, vol. 20, no. 21, p. 6281, Nov. 2020.

[21] J. Jia, M. Fu, X. Liu, and B. Zheng, "Underwater object detection based on improved EfficientDet," *Remote Sens.*, vol. 14, no. 18, p. 4487, Sep. 2022.

[22] F. Han, J. Yao, H. Zhu, and C. Wang, "Underwater image processing and object detection based on deep CNN method," *J. Sensors*, vol. 2020, pp. 1–20, May 2020.

[23] B. Xue, B. Huang, W. Wei, G. Chen, H. Li, N. Zhao, and H. Zhang, "An efficient deep-sea debris detection method using deep neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 12348–12360, 2021.

[24] C.-H. Yeh, C.-H. Lin, L.-W. Kang, C.-H. Huang, M.-H. Lin, C.-Y. Chang, and C.-C. Wang, "Lightweight deep neural network for joint learning of underwater object detection and color conversion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 11, pp. 6129–6143, Nov. 2022.

[25] C. Liu, Z. Wang, S. Wang, T. Tang, Y. Tao, C. Yang, H. Li, X. Liu, and X. Fan, "A new dataset, Poisson GAN and AquaNet for underwater object grabbing," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 5, pp. 2831–2844, May 2022.

[26] K. Panetta, L. Kezebou, V. Oludare, and S. Agaian, "Comprehensive underwater object tracking benchmark dataset and underwater image enhancement with GAN," *IEEE J. Ocean. Eng.*, vol. 47, no. 1, pp. 59–75, Jan. 2022.

[27] J. Liu, S. Liu, S. Xu, and C. Zhou, "Two-stage underwater object detection network using Swin transformer," *IEEE Access*, vol. 10, pp. 117235–117247, 2022.

[28] N. A. Ubina, S.-C. Cheng, C.-C. Chang, S.-Y. Cai, H.-Y. Lan, and H.-Y. Lu, "Intelligent underwater stereo camera design for fish metric estimation using reliable object matching," *IEEE Access*, vol. 10, pp. 74605–74619, 2022.

[29] D. Sierra-Sosa, M. Telahun, and A. Elmaghraby, "TensorFlow quantum: Impacts of quantum state preparation on quantum machine learning performance," *IEEE Access*, vol. 8, pp. 215246–215255, 2020.

[30] T. M. Khan and A. Robles-Kelly, "Machine learning: Quantum vs classical," *IEEE Access*, vol. 8, pp. 219275–219294, 2020.

[31] P. Easom-McCaldin, A. Bouridane, A. Belatreche, R. Jiang, and S. Al-Maadeed, "Efficient quantum image classification using single Qubit encoding," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 17, 2022, doi: 10.1109/TNNLS.2022.3179354.

[32] X. Zhou, Y. Feng, and S. Li, "Supervised learning enhanced quantum circuit transformation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 42, no. 2, pp. 437–447, Feb. 2023, doi: 10.1109/TCAD.2022.3179223.

[33] I. Nikoloska and O. Simeone, "Training hybrid classical-quantum classifiers via stochastic variational optimization," *IEEE Signal Process. Lett.*, vol. 29, pp. 977–981, 2022.

[34] S. Otgonbaatar and M. Datcu, "Classification of remote sensing images with parameterized quantum gates," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[35] A. Burguera, "Segmentation through patch classification: A neural network approach to detect *Posidonia oceanica* in underwater images," *Ecological Informat.*, vol. 56, Mar. 2020, Art. no. 101053.

[36] P. X. Huang, B. J. Boom, and R. B. Fisher, "Hierarchical classification with reject option for live fish recognition," *Mach. Vis. Appl.*, vol. 26, no. 1, pp. 89–102, Jan. 2015.

[37] N. L. Chau, T.-P. Dao, and V. T. T. Nguyen, "An efficient hybrid approach of finite element method, artificial neural network-based multiobjective genetic algorithm for computational optimization of a linear compliant mechanism of nanoindentation tester," *Math. Problems Eng.*, vol. 2018, pp. 1–19, Oct. 2018.

[38] G. R. Ríos-Muñoz, F. Fernández-Avilés, and Á. Arenal, "Convolutional neural networks for mechanistic driver detection in atrial fibrillation," *Int. J. Mol. Sci.*, vol. 23, no. 8, p. 4216, Apr. 2022.

[39] S. Villon, D. Mouillot, M. Chaumont, E. S. Darling, G. Subsol, T. Claverie, and S. Villéger, "A deep learning method for accurate and fast identification of coral reef fishes in underwater images," *Ecological Informat.*, vol. 48, pp. 238–244, Nov. 2018.

[40] A. B. Labao and P. C. Naval, "Cascaded deep network systems with linked ensemble components for underwater fish detection in the wild," *Ecological Informat.*, vol. 52, pp. 103–121, Jul. 2019.

[41] X. Guo, X. Zhao, Y. Liu, and D. Li, "Underwater sea cucumber identification via deep residual networks," *Inf. Process. Agricult.*, vol. 6, no. 3, pp. 307–315, Sep. 2019.

[42] E. Prasetyo, N. Suciati, and C. Fatichah, "Multi-level residual network VGGNet for fish species classification," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 8, pp. 5286–5295, Sep. 2022.

[43] A. Gómez-Ríos, S. Tabik, J. Luengo, A. S. M. Shihavuddin, and F. Herrera, "Coral species identification with texture or structure images using a two-level classifier based on convolutional neural networks," *Knowl.-Based Syst.*, vol. 184, Nov. 2019, Art. no. 104891.

[44] X. Wang, J. Jiao, J. Yin, W. Zhao, X. Han, and B. Sun, "Underwater sonar image classification using adaptive weights convolutional neural network," *Appl. Acoust.*, vol. 146, pp. 145–154, Mar. 2019.

[45] K. Zhang, W. Wang, Z. Lv, Y. Fan, and Y. Song, "Computer vision detection of foreign objects in coal processing using attention CNN," *Eng. Appl. Artif. Intell.*, vol. 102, Jun. 2021, Art. no. 104242.

[46] N. F. F. Alshdaifat, A. Z. Talib, and M. A. Osman, "Improved deep learning framework for fish segmentation in underwater videos," *Ecological Informat.*, vol. 59, Sep. 2020, Art. no. 101121.

[47] A. Ganesan and S. M. Santhanam, "A novel feature descriptor based coral image classification using extreme learning machine with ameliorated chimp optimization algorithm," *Ecological Informat.*, vol. 68, May 2022, Art. no. 101527.

[48] J. Zhou, T. Yang, and W. Zhang, "Underwater vision enhancement technologies: A comprehensive review, challenges, and recent trends," *Int. J. Speech Technol.*, vol. 53, no. 3, pp. 3594–3621, Feb. 2023.

[49] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[50] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[51] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2818–2826.

[52] D. Xue, X. Zhou, C. Li, Y. Yao, M. M. Rahaman, J. Zhang, H. Chen, J. Zhang, S. Qi, and H. Sun, "An application of transfer learning and ensemble learning techniques for cervical histopathology image classification," *IEEE Access*, vol. 8, pp. 104603–104618, 2020.