Real-time Hybrid Facial Expression Analysis Model Using Quantum Distance-based Classifier and Classical Artificial Neural Networks : QUADCANN

Karthikeyan Rengasamy

karthikeyan.r@iiits.in

Indian Institute of Information Technology Sri City

Piyush Joshi

Indian Institute of Information Technology Sri City

Raveendra VVS

Tata Consultancy Services (India)

Research Article

Keywords: Quantum Facial Expression Analysis, Quantum Image Classification, Quantum-Classical Hybrid Model, Quantum Distance-based Classifier, Hybrid ANN FEA

Posted Date: January 11th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3850071/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Quantum Information Processing on June 1st, 2024. See the published version at https://doi.org/10.1007/s11128-024-04437-3.

Real-time Hybrid Facial Expression Analysis Model Using Quantum Distance-based Classifier and Classical Artificial Neural Networks : QUADCANN

Karthikeyan Rengasamy · Piyush Joshi · Raveendra VVS

Received: date / Accepted: date

Abstract Facial expressions bring human interactions to life with nonverbal cues that convey hues of emotions, feelings, and cultural intent. Analyzing these expressions is essential in the age of digital transformation. Traditional approaches to real-time facial expression analysis have limitations in capturing the complexity of these expressions and consume enormous computational power. The promising computing capacity of quantum computers is poised to solve current problems and set new standards by meeting growing demands. We have created a hybrid model that leverages the capabilities of the classical and quantum domains. The proposed model uses a quantum distance-based classifier along with classical artificial neural networks to perform real-time facial expression analysis and classification. Based on a thorough review of the literature, this novel and comprehensive work is, to the best of our knowledge, one of the few that addresses eight different emotions in the quantum domain. The implementation of a novel quantum error correction method has improved the accuracy of this hybrid model. Our model was trained on the CK+ facial expression database. To ensure fairness of the study, we tested our model on the largest facial expression database, AffectNet-8, and compared the performance with state-of-the-art models. Our model accuracy is 10.83% higher than that of the state-of-the-art models. From a holistic perspective, our proposed novel hybrid model appears to have universal value for all kinds of real-time image analysis and classification problems. As we move forward, we planned to focus on quantum neural networks for image processing and facial expression analysis.

Keywords Quantum Facial Expression Analysis · Quantum Image Classification · Quantum-Classical Hybrid Model · Quantum Distance-based Classifier · Hybrid ANN FEA

Sponsored by TATA Consultancy Services

Indian Institute of Information Technology Sri City
Chittoor, India
E-mail: karthikeyan.r@iiits.in

Indian Institute of Information Technology Sri City
Chittoor, India
E-mail: piyush.joshi@iiits.in

TATA Consultancy Services
Chennai, India
E-mail: vvs.raveendra@tcs.com

1 Introduction

The emotional aspects of facial expressions are crucial for pragmatic study, interpretation and deeper assimilation of human behaviour. Traditional facial expression analysis (FEA) models typically rely on machine learning (ML) algorithms [1] running on classical computers. However, the ever-growing demand for faster and more efficient real-time image processing [2] has spurred interest in exploring the potential of novel approaches such as quantum computing [3], to accelerate FEA tasks.

In this work, we present QUADCANN, a novel hybrid FEA model, which combines classical artificial neural networks (ANN) [4] with a quantum distance-based classifier (QDC) [5]. Our proposed hybrid model leverages the strengths of both deep learning (DL) and quantum computing [6] to enhance the accuracy and performance of facial expression classification tasks [7]. The distances between test instances and training instances were measured by

the QDC [5]. The measured distances and its labels were used to train the ANN model [4]. Subsequently, we account for error handling [8, 9] and normalize the output to improve the model's accuracy. In this paper, we considered 8 facial expressions: {Angry (A), Contempt (C), Disgust (D), Fear (F), Happy (H), Neutral (N), Sad (S), and Surprise (O)} for classification.

The rest of the paper is organized as follows. Recent works in FEA [5, 10–12] and QDC [4, 9] research are summarized in Section 2. Specifics of QUADCANN model, QDC, ANN architecture, and novel error correction method are detailed in Section 3. Our computational experiment setup and findings from the comparative study are presented in Section 4. Section 5 concludes with our conclusions and recommendations for further research.

2 Related Work

This section presents recent studies on image classification, encompassing facial recognition, feature extraction, and emotion detection. With the rise of neural networks in computer vision, many FEA models combined with neural networks have emerged. Razuri et al. [10] developed an automated emotion recognition system using an ANN. In this study, neural networks are static networks that do not have any temporal relationships between inputs and outputs. Nath et al. [11] presented a comprehensive survey of image classification techniques, and Giuseppe et al. [12] compared face verification methods. Owayjan et al. [4] used a multilayer perceptron with backpropagation in the development of an ANN model for feature extraction and classification. Only three expressions can be classified by the model: neutral, angry, and happy. Nonetheless, these results are highly precise and have opened the door for the further development of real-time FEA applications.

Fathallah et al. [13]. presented a novel convolutional neural network (CNN) architecture for FEA. The network was refined using the visual geometry group (VGG) model to achieve better accuracy. The results demonstrate that the refined CNN approach shows improvements in facial expression analysis. The model has also been tested on numerous public databases and shows consistent results. Jain et al. [14] developed a deep CNN model that contains two residual blocks, each containing four convolutional layers. The model was trained on Japanese female facial expression (JAFFE) [15] and extended Cohn-Kanade CK+ [16] databases following a preprocessing step that allows image cropping and intensity normalization. Yue et al. [17] used a modular multichannel deep convolutional neural network to construct an FEA model. A global average layer was used in the network output to prevent overfitting. Data enhancement on the dataset prior to training increased the model's capacity for generalization.

As stated by Christopher et al. [18], due to computational constraints, the majority of the classical models are trained using low resolution grayscale image datasets. To obtain the required accuracy, a large dataset of training images and high processing power are needed. Quantum computing (QC) has the potential to revolutionize image processing and offers a new approach to solving these problems. QC can execute certain complex image processing ML algorithms much faster and more efficiently [19] than classical computers.

Schuld et al. [5] built a quantum interference circuit using Hadamard gates to compute the distance between the image features in a quantum parallel and explored the potential of the QDC. The model by Mengoni et al. [20] utilizes Schuld et al.'s [5] quantum interference circuit for classifying facial expressions. However, Mengoni et al. [20] identified the closest match by evaluating the distance between a test instance and a random instance from each training batch. Hence, representative selection becomes crucial for the model. Khurelsukh et. al. [21] attempted to solve the data re-uploading concept of the FEA problem on the qubit(s) of parameterized quantum circuits in the model's convolutional layer. This resulted in a validation accuracy of 70.25% on the test set. Subsequently, an ensemble model averaging method was used, and the results encouraged further research. Bhatt et al. [22] created a model by applying meta heuristic algorithm inspired by quantum mechanics. The deep CNN receives features and adds residual blocks for classification. As the yaw angle increases, the algorithm performance decreases.

Several types of noise can affect the output of quantum computing. Therefore, it is important to consider the different types of quantum error to achieve near-error-free output. Majumdar et al. [23] offer an error estimation technique through sparse scheduling, that can help lower the quantity of error correcting blocks needed. Swathi et al. [24] proposed a qubit encoding method for the correction of quantum errors. Kukulski et al. [8] used probabilistic error correction methods to correct the measurement errors. The probability of successful error correction required an additional physical qubit to restore a qubit state.

Rengasamy et al. [25], aimed to solve the FEA problem by combining QDC with a classical linear

support vector machine in a hybrid model. The model reached its maximum accuracy of 85% against the custom-built FEA dataset, leaving scope for further improvement. Zheng et al. [26] created a two-stream pyramid cross-fusion transformer network (POSTER) model to merge image features with facial landmark features. However, POSTER's enormous computational effort has led researchers to look for ways to further improve the model. Mao et al. [27] sought to resolve this problem. To achieve the required accuracy, the architecture of POSTER was enhanced and POSTER++ was proposed, which combines images with multiscale features of landmarks. Nevertheless, the model achieved a maximum accuracy of 63.77% on the AffectNet-8 [28] dataset, which leaves room for further improvement.

The QUADCANN model draws inspiration from Owayjan et al.'s [4] use of artificial neural networks, Mengoni et al.'s [20] QDC, Kukulski et al.'s [8] quantum error correction method and Rengasamy et al.'s [25] hybrid approach to further increase the accuracy.

3 Quantum-Classical Hybrid Model

Current research shows that there is a growing interest in combining different models [11]. Aligned with the current trend, we propose a quantum classical hybrid classifier model, as shown in Fig. 1. This model combines the computational power of QC with the accuracy of the classical DL algorithm for better prediction. In a classical computer, an image is constructed as an array of pixels with corresponding RGB color intensities and locations. It is merely a matrix representation of classical bits. On a quantum computer, all types of images or data are represented as qubits, denoting the quantum states that hold the data. Image analysis and conversion into a quantum state representation [29] in qubits are performed before processing the image using a quantum algorithm. We have briefly illustrated the proposed work in the following subsections.

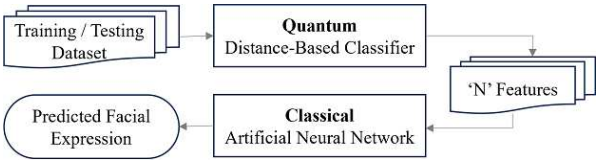


Fig. 1. Quantum-Classical Hybrid Model

3.1 Proposed Algorithm

Our algorithm considers the FEA techniques described in the introduction, as shown in Fig. 2. We begin by identifying facial landmarks using the 68-point landmark detection method. Following the detection of landmark points, the Delaunay triangulation [30] method is used to extract features from mouth points, which are represented as eigenvectors and values in a hollow symmetric matrix. The distances between each training and test instances are then computed using the QDC and normalized, accounting for quantum measurement errors. The training set of an ANN is constructed with these normalized distances and used for classification prior to putting the model under test. After training, the model is used to classify the given facial expression.

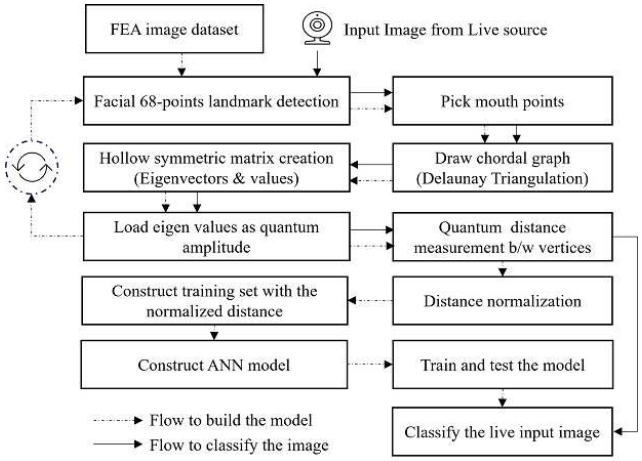


Fig. 2. Quantum-Classical Hybrid Model Algorithm

3.2 FEA Image Dataset

The CK+ [16] database is regarded as most widely used facial expression database in FEA models. QUADCANN model also required high resolution RGB and grayscale images to study efficiency. Using images with the resolution of 640×480 from CK+ database, we were able to generate high resolution (2048×1080) image dataset. Therefore, the CK+ database was used to train the QUADCANN model. The AffectNet database [28] was used to benchmark the efficiency of SOTA models. For fair comparison, the AffectNet database was used to test the efficiency of the QUADCANN model.

3.3 Image Segmentation using 68-points Landmark Detection

Detection of facial landmarks forms the basis for FEA, and 68-point landmarks on a human face are detected using the histogram of oriented gradients (HOG) DL method [12], as shown in Fig. 3.

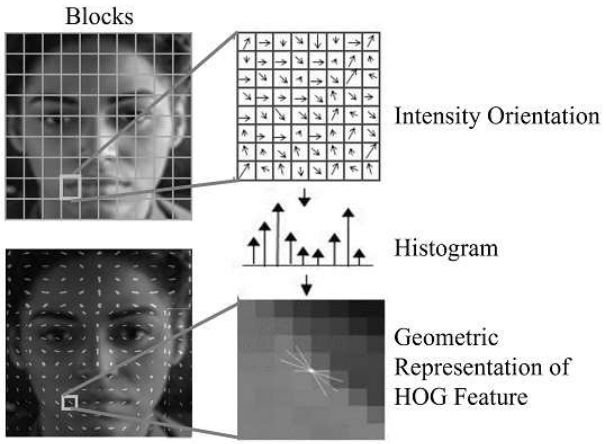
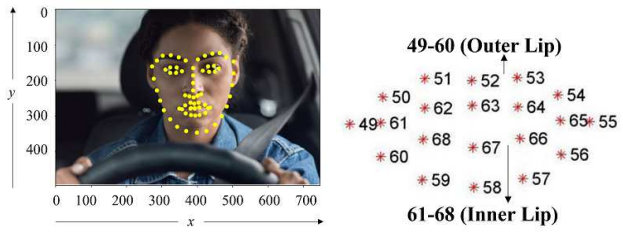


Fig. 3. Histogram of Orientated Gradient Extraction from Face

These points are represented as an array of 68 (x, y) coordinates, with $x, y \in \mathbb{R}$ mapped to the facial structures, as shown in Fig. 4a. Features of lips are the anchors for all known facial expressions. For lip feature extraction, the outer (49-60) and inner (61-68) lip points are selected from the 68-point facial landmarks, as shown in Fig. 4b.



(a) (b)

Fig. 4. (a) 68 Landmark Points Marked (b) Region of Interest from 68 Landmark Points

After successfully detecting these inner and outer lip landmark points, feature extraction was performed.

3.4 Feature Extraction

Feature extraction creates a weighted undirected graph using Delaunay Triangulation [30] and chordal graph [31] methods. The resulting chordal graph is represented by an $N \times N$ matrix, where N is the number of vertices. Each vertex is represented in the form of (x, y), as shown in Fig. 5.

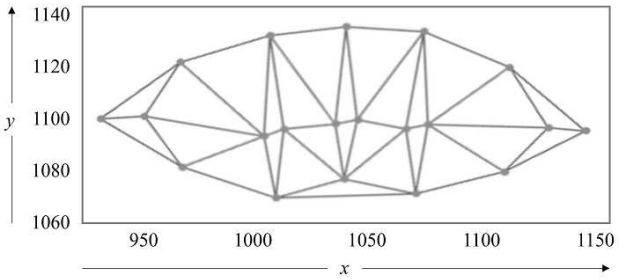


Fig. 5. Chordal Graph Encoding using Delaunay Triangulation

After feature extraction, image classification is performed using the QDC algorithm. In addition to the QDC, we conducted a comparative analysis with the classical distance-based classifier (CDC).

3.5 Quantum Distance-based Classifier(QDC)

A QDC is constructed based on quantum computing principles such as superposition, entanglement, and interference to calculate the distance between data points. This involves encoding the data into quantum states, using quantum algorithms to measure the distances and classify the data. Ten quantum registers were used to build the QDC circuit [9] for the specific FEA problem addressed in this context, as shown in Fig. 6. This circuit was constructed using auxiliary $[|a\rangle]$, index $[|i\rangle]$, data $[|d\rangle]$, and class $[|c\rangle]$ registers.

In quantum computing, for a feature vector $G = [g_1, g_2, ..., g_d]$ represents the chordal graph, and the corresponding quantum state of the vector $|G\rangle$ is represented by elements d and normalization constant γ, as in Eq. (1).

$$|G\rangle = \frac{1}{\gamma} \sum_{k=1}^{d} g_k |k\rangle \quad where\ \gamma = \sqrt{\sum_{k=1}^{d} |g_k|^2} \tag{1}$$

The distance between the vectors $(qd_{test \to x})$ is calculated using quantum interference with G_{test}

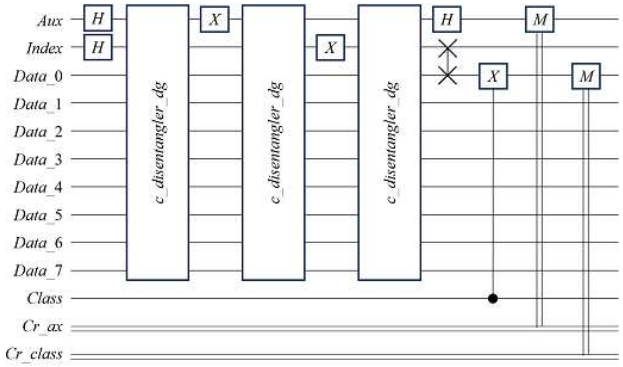


Fig. 6. Quantum Distance-based Classifier Circuit

(vector of test face) and G_x (vector of 8 expression faces), as in Eq. (2).

$$\begin{aligned} qd_{test\to x} &= QuantumDistance(|G_{test}\rangle, |G_x\rangle) \\ &= \frac{1}{2\sqrt{2}}(|0\rangle_a(|G_{test}\rangle + |G_x\rangle)_d) \\ &\quad + (|1\rangle_a(|G_{test}\rangle - |G_x\rangle)_d)|x\rangle_i|y_x\rangle_c \\ &= \frac{1}{2\sqrt{2}} \sum_{Q=|0\rangle}^{|1\rangle} (Q_a(|G_{test}\rangle + |G_x\rangle)_d)|x\rangle_i|y_x\rangle_c \end{aligned} \tag{2}$$

The accuracy of the QDC is determined solely by the representatives of each emotion in the training set. This made us think about making further improvements to achieve better accuracy. To increase the accuracy, we combined the QDC model with an ANN, as it has the ability to learn from input data. After training, the ANN can predict the right output for the given input that it has never "seen" before.

3.6 Quantum Error Correction (QEC)

Quantum computers are vulnerable to multiple sources of noise, such as qubit decoherence, individual gate errors, and measurement errors. During measurement, errors are added to the qubits. Due to quantum measurement errors [32], it is possible that a class receives greater weight. Therefore, QDC predictions are prone to bias. To address these errors, we propose to modify Kukulski et al's [8] probabilistic error correction method, as shown in Fig. 7. This was achieved by using the auxiliary qubit as a probabilistic error checker. The auxiliary qubit was set to either $|0\rangle$ or $|1\rangle$ before executing the circuit. After execution, output of the auxiliary qubit was compared to it's original state. If the final state of the auxiliary qubit differed from its initial state, the output was deemed to be an error. As a result, the overall output does not include this measurement and was excluded. The output was regarded as expected only if the initial and final states of the auxiliary qubit were measured and found to be the same.

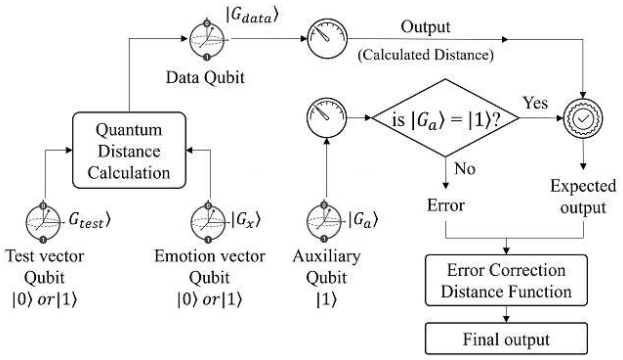


Fig. 7. Modified probabilistic error-correcting procedure

To accommodate this QEC method, we propose a change in the distance function, as in Eq. (3). To the best of our knowledge, this novel quantum error correction method is the first of its kind for real-time image processing.

$$\begin{aligned} QEC &= D(|test\rangle, |x\rangle) = \varepsilon_{corr}(qd_{test\to x}) \\ &= N \times \frac{qd_{test\to x}}{\sum_{v\in[X]} P(|G_v\rangle)} \\ &= \frac{N}{2\sqrt{2}} \times \frac{\sum_{Q=|0\rangle}^{|1\rangle} (Q_a(|G_{test}\rangle + |G_x\rangle)_d)|x\rangle_i|y_x\rangle_c}{\sum_{v\in[X]} P(|G_v\rangle)} \end{aligned} \tag{3}$$

where $[X] = \{A, C, D, F, H, N, S, O\}$. To enhance accuracy, unlike classical models, N shots were taken with the same representatives, and probabilities $P(|G_v\rangle)$ were calculated. Adjusting for errors and normalized distances, a training set is then created by transforming the symmetric $N \times N$ matrix into eigenvectors and eigenvalues.

3.7 Creation of Training Set for ANN

Training sets for all eight emotions are created using the distance measured by the QDC, as in Eq. 4, as shown in Fig. 8.

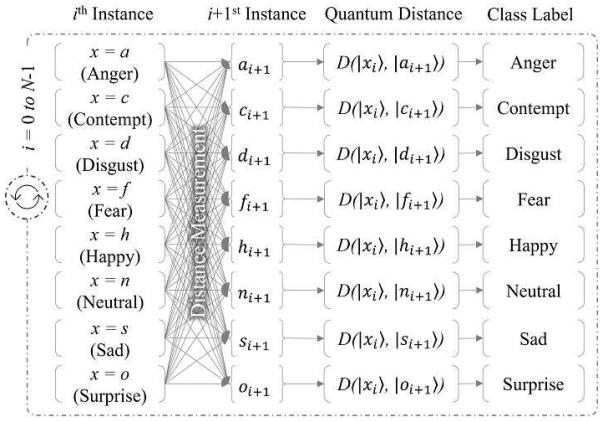


Fig. 8. Training Set Creation Using the Distances Measured by the QDC as Features

$$S_X = \bigcup_{i=1}^{m-1} [D(|x_i\rangle, |a_{i+1}\rangle), D(|x_i\rangle, |c_{i+1}\rangle), D(|x_i\rangle, |d_{i+1}\rangle), D(|x_i\rangle, |f_{i+1}\rangle), D(|x_i\rangle, |h_{i+1}\rangle), D(|x_i\rangle, |n_{i+1}\rangle), D(|x_i\rangle, |s_{i+1}\rangle), D(|x_i\rangle, |o_{i+1}\rangle), Y_i] \quad (4)$$

where X is the test emotion instance vector, x is the eigenvalue of the emotion instance vector X, a, c, d, f, h, n, s, and o are the eigenvalues of the emotion instance vectors, m is the number of instances, and Y is an actual class of the vector. To build the QUADCANN model, the datasets generated above are then merged together, as in Eq. 5.

$$S = \bigcup_{v\epsilon[X]} S_v = S_A \cup S_C \cup S_D \cup S_F \cup S_H \cup S_N \cup S_S \cup S_O \quad (5)$$

Table I shows the complete training set generated from the Eq. 5.

3.8 Feed Forward to Classical ANN

Multiclass classification of the nonlinear inputs is performed via the ANN architecture, as shown in Fig. 9. A network is constructed with the input layer accepting eight input features and two hidden layers each having 25 neurons with ReLU [33] activation functions, as in Eqs. 6 & 7.

$$z = \sum_{i=0}^{N-1} W_i F_x + b = W_0F_a + W_1F_c + W_2F_d + W_3F_f + W_4F_h + W_5F_n + W_6F_s + W_7F_o + b \quad (6)$$

where N = The number of input features, W_i = The weights calculated using gradient descent [34], $F_{x\epsilon[a,c,d,f,h,n,s,o]} = D_{x\epsilon[a,c,d,f,h,n,s,o]}$ (distance to respective emotion instance vectors), and b = bias.

$$R(z) = max(0, z) = Z_j \quad (7)$$

A softmax [35] function is employed in the classifier's output layer to provide class decision probabilities, as shown in Eq. 8.

$$\text{Softmax}(Z_j) = \sigma(Z_j) = \frac{e^{Z_j}}{\sum_{i=1}^{N} e^{Z_i}} \quad (8)$$

where Z_j = Input of the softmax function.

During training, the network employs a backpropagation technique to minimize the binary cross-entropy loss function, as shown in Eq. 9. This network calculates the weights of all classes.

$$H(Y_j, \hat{Y_j}) = -\sum_{j\epsilon[a,c,d,f,h,n,s,o]} Y_j log\hat{Y_j} \quad (9)$$

where $\hat{Y_j}$ is the predicted label, which is Softmax(Z_j) and Y_j is the true label. After training the model, predictions were made based on decision probabilities. The model was observed to be consistent with its informed prediction by churning out either an extremely positive or negative output even in an imperfect situation.

4 Results Analysis

Along with the construction and assessment of the QUADCANN model, we studied the performance of the CDC and QDC models in isolation. Subsequently, the results of these studies were compared and interpreted.

4.1 Results on the CK+ dataset

The QUADCANN model was trained using a nonlinear dataset (a subset of the CK+ dataset [16]) as shown in Fig. 10. For comparison, grayscale and RGB images comprising low-resolution (LD) (96×96) and high-resolution (HD) (1024×1024) images were used. We used both training and testing accuracy as the basis for our evaluation. Note that there would be minor variations in the model's accuracy between its independent training phase and the testing phase.

TABLE I. Training set generated from the QDC for ANN classifier with normalized distances as features $(Fa, Fc, Fd, Ff, Fh, Fn, Fs, Fo)$ of respective emotions instance vector

Fa[1]	Fc[2]	Fd[3]	Ff[4]	Fh[5]	Fn[6]	Fs[7]	Fo[8]	Y[9]
0	$D(\|a_i\rangle, \|c_{i+1}\rangle)$	$D(\|a_i\rangle, \|d_{i+1}\rangle)$	$D(\|a_i\rangle, \|f_{i+1}\rangle)$	$D(\|a_i\rangle, \|h_{i+1}\rangle)$	$D(\|a_i\rangle, \|n_{i+1}\rangle)$	$D(\|a_i\rangle, \|s_{i+1}\rangle)$	$D(\|a_i\rangle, \|o_{i+1}\rangle)$	Anger
0	...	...	...	...	...	...	...	Anger
$D(\|c_i\rangle, \|a_{i+1}\rangle)$	0	$D(\|c_i\rangle, \|d_{i+1}\rangle)$	$D(\|c_i\rangle, \|f_{i+1}\rangle)$	$D(\|c_i\rangle, \|h_{i+1}\rangle)$	$D(\|c_i\rangle, \|n_{i+1}\rangle)$	$D(\|c_i\rangle, \|s_{i+1}\rangle)$	$D(\|c_i\rangle, \|o_{i+1}\rangle)$	Contempt
...	0	...	...	...	...	...	...	Contempt
$D(\|d_i\rangle, \|a_{i+1}\rangle)$	$D(\|d_i\rangle, \|c_{i+1}\rangle)$	0	$D(\|d_i\rangle, \|f_{i+1}\rangle)$	$D(\|d_i\rangle, \|h_{i+1}\rangle)$	$D(\|d_i\rangle, \|n_{i+1}\rangle)$	$D(\|d_i\rangle, \|s_{i+1}\rangle)$	$D(\|d_i\rangle, \|o_{i+1}\rangle)$	Disgust
...	...	0	...	...	...	...	...	Disgust
$D(\|f_i\rangle, \|a_{i+1}\rangle)$	$D(\|f_i\rangle, \|c_{i+1}\rangle)$	$D(\|f_i\rangle, \|d_{i+1}\rangle)$	0	$D(\|f_i\rangle, \|h_{i+1}\rangle)$	$D(\|f_i\rangle, \|n_{i+1}\rangle)$	$D(\|f_i\rangle, \|s_{i+1}\rangle)$	$D(\|f_i\rangle, \|o_{i+1}\rangle)$	Fear
...	...	...	0	...	...	...	...	Fear
$D(\|h_i\rangle, \|a_{i+1}\rangle)$	$D(\|h_i\rangle, \|c_{i+1}\rangle)$	$D(\|h_i\rangle, \|d_{i+1}\rangle)$	$D(\|h_i\rangle, \|f_{i+1}\rangle)$	0	$D(\|h_i\rangle, \|n_{i+1}\rangle)$	$D(\|h_i\rangle, \|s_{i+1}\rangle)$	$D(\|h_i\rangle, \|o_{i+1}\rangle)$	Happy
...	...	...	...	0	...	...	...	Happy
$D(\|n_i\rangle, \|a_{i+1}\rangle)$	$D(\|n_i\rangle, \|c_{i+1}\rangle)$	$D(\|n_i\rangle, \|d_{i+1}\rangle)$	$D(\|n_i\rangle, \|f_{i+1}\rangle)$	$D(\|n_i\rangle, \|h_{i+1}\rangle)$	0	$D(\|n_i\rangle, \|s_{i+1}\rangle)$	$D(\|n_i\rangle, \|o_{i+1}\rangle)$	Neutral
...	...	...	...	...	0	...	...	Neutral
$D(\|s_i\rangle, \|a_{i+1}\rangle)$	$D(\|s_i\rangle, \|c_{i+1}\rangle)$	$D(\|s_i\rangle, \|d_{i+1}\rangle)$	$D(\|s_i\rangle, \|f_{i+1}\rangle)$	$D(\|s_i\rangle, \|h_{i+1}\rangle)$	$D(\|s_i\rangle, \|n_{i+1}\rangle)$	0	$D(\|s_i\rangle, \|o_{i+1}\rangle)$	Sad
...	...	...	...	...	...	0	...	Sad
$D(\|o_i\rangle, \|a_{i+1}\rangle)$	$D(\|o_i\rangle, \|c_{i+1}\rangle)$	$D(\|o_i\rangle, \|d_{i+1}\rangle)$	$D(\|o_i\rangle, \|f_{i+1}\rangle)$	$D(\|o_i\rangle, \|h_{i+1}\rangle)$	$D(\|o_i\rangle, \|n_{i+1}\rangle)$	$D(\|o_i\rangle, \|s_{i+1}\rangle)$	0	Surprise
...	...	...	...	...	...	...	0	Surprise

[1] Distance to Anger, [2] Distance to Contempt, [3] Distance to Disgust, [4] Distance to Fear, [5] Distance to Happy, [6] Distance to Neutral, [7] Distance to Sad, [8] Distance to Surprise, [9] Actual Class Label

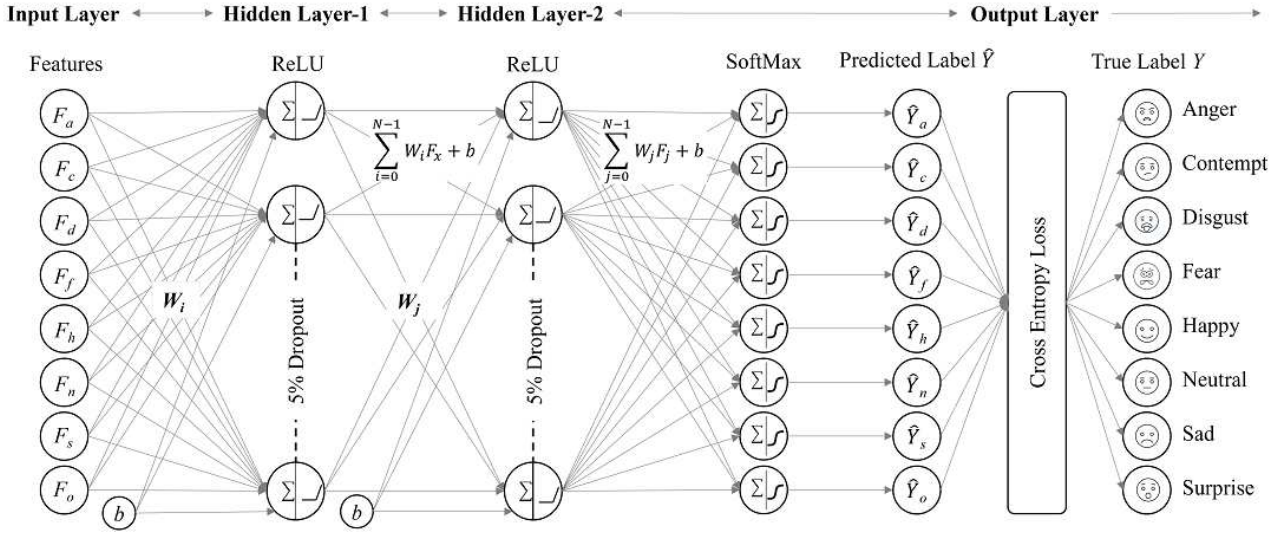


Fig. 9. ANN Architecture for FEA

4.1.1 Model Accuracy

The accuracy of CDC and QDC [9] in isolation is in the range of 35-50% for all types of image datasets, indicating scope for further improvement. With a combination of CDC and ANN [5] models, we observed a significant improvement in accuracy (85%) for high-resolution RGB images. However, we deemed this model to be underfitting, as it performed very well on training data, albeit poorly on all types of test image dataset. We attributed this to the fact that the extraction of the CDC features was highly consistent, which may not be the case in reality. After incorporating the QEC, the QUADCANN model achieved a maximum accuracy of 95% for high-resolution RGB image datasets, as shown in Fig. 11.

4.1.2 Accuracy Comparison

For high-resolution grayscale image datasets, the QUADCANN model achieved an accuracy of 75%, which is 25% greater than that of distance-based classifiers on both classical and quantum methods. QUADCANN has 2% greater accuracy than the classical counterpart model, CDC-ANN. The QUADCANN model achieved 95% accuracy on high resolution RGB image datasets, which is 10% better than that of its classical counterpart, CDC-ANN. QUADCANN achieved at least 20%

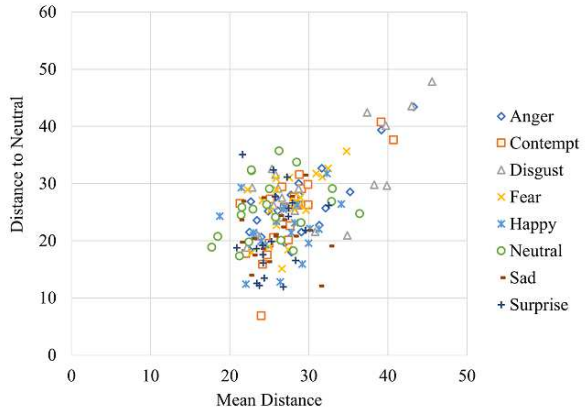


Fig. 10. Custom Training dataset (subset of the CK+ dataset)

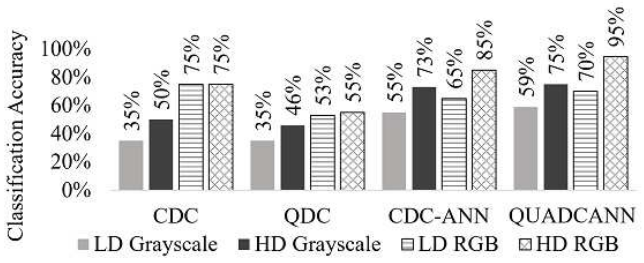


Fig. 11. Accuracy Comparison between CDC, QDC and Hybrid Models

higher accuracy when compared to pure distance-based classifiers. These observations reiterate the importance of utilizing error correction methods while developing quantum models. Furthermore, the model achieved significantly higher accuracy with fewer high-resolution RGB images, as shown in Fig. 12.

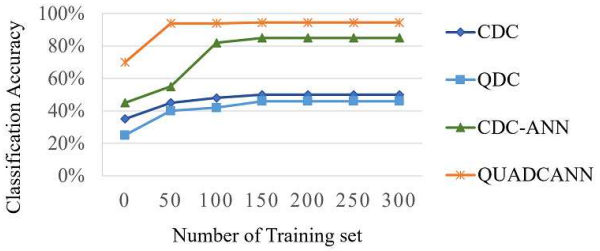


Fig. 12. Accuracy w.r.t. High-resolution RGB Image Dataset

Although high-resolution RGB images delivered better accuracy in the ANN models, grayscale images were consistent across all models. Because it is primarily in 2 spatial dimensions (2D) with 1 layer (Gray) rather than 2D with 3 layer (RGB), it enables easy differentiation of shadow details and highlights of given image features.

4.1.3 Loss vs Accuracy

As the initial weights of the QUADCANN model were assigned at random, it experienced a high initial average loss. However, as the model progresses through the training process, the average loss gradually decreases and reaches a saturation point after 800 epochs, as shown in Fig. 13a.

Compared to classical models, the QUADCANN model requires fewer training datasets. As a result, the accuracy of the QUADCANN model reached its peak in approximately 200 epochs, as shown in Fig. 13b. We attribute this to the deployment of QEC to optimize the features extracted from QDC. This approach enabled the QUADCANN model to identify the optimal classification parameters and achieve maximum accuracy in a relatively small number of epochs. The results also highlight the potential benefits of incorporating QEC techniques into QC models. To avoid overfitting, early stopping was implemented at approximately 300 epochs when parameter optimization no longer resulted in a significant increase in accuracy, as shown in Fig. 13b.

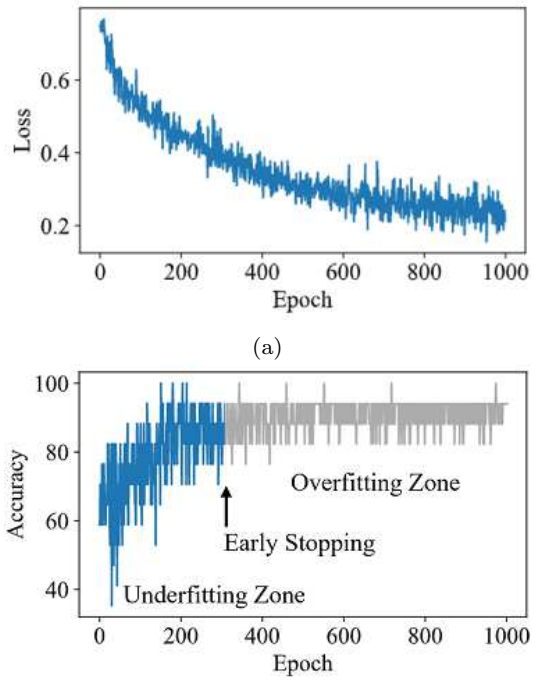
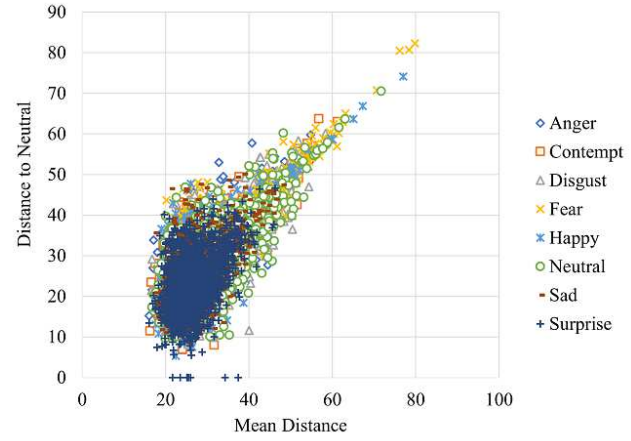


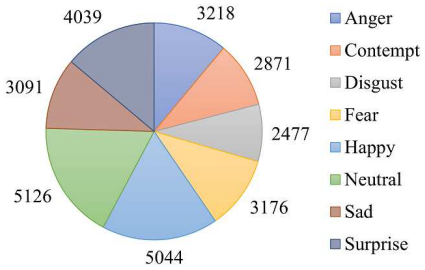
Fig. 13. (a) Loss w.r.t. Epoch (b) Training Accuracy w.r.t. Epoch

4.2 Results on the AffectNet-8 dataset

To ensure fairness in the study and for comparison with the SOTA models, QUADCANN model was also tested using the AffectNet-8 dataset, as shown in Figs. 14a, 14b & 14c. AffectNet is a sizable FEA dataset with 0.4 million images that have been meticulously tagged with the presence of eight different emotions, as mentioned in Section 1, along with their corresponding valence and intensity of arousal. The QUADCANN model is capable of classifying complex nonlinear dataset effectively and eventually achieved greater accuracy.



(a)



(b)

(c)

Fig. 14. (a) AffectNet-8 dataset (b) Class Distribution (c) Sample Images

4.2.1 Model Accuracy

With respect to the images in the AffectNet-8 dataset, the QUADCANN model achieved a maximum accuracy of 74.6% at 200 epochs and learning rate of .001. Hence, we stop training the model at 200 epochs to reduce computational costs and additional complexities.

4.2.2 Accuracy Comparison with SOTA FEA Models

We performed a study comparing the accuracy of our QUADCANN model with that of other SOTA models, using the AffectNet-8 dataset. The accuracy of all the SOTA models considered in our study was restricted to the range of 61.09% to 63.77%. Compared to the most recent model, QUADCANN demonstrated a higher accuracy of 74.60%, a 10.83% increase in accuracy, as shown in Fig. 15. This is because the features produced by the classifier based on quantum distance are superior to those extracted directly from the images. We have taken into consideration the two models from the top of the table, POSTER and POSTER++, for additional comparisons at the emotional level.

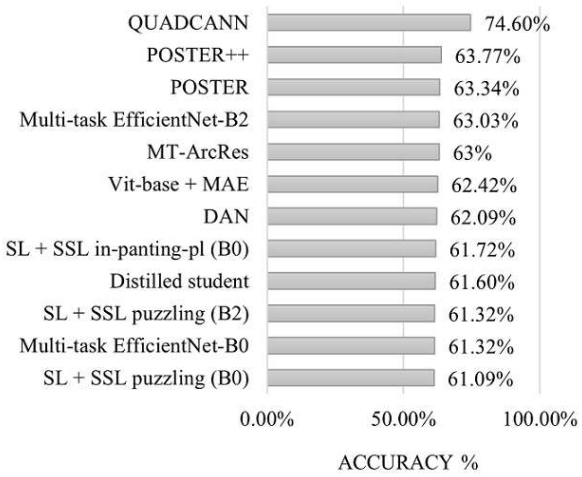


Fig. 15. Comparison of Accuracy with State-of-the-art Models using the AffectNet-8 dataset

4.2.3 Accuracy Comparison at the Emotional Level

Subsequently, we performed a study at the emotional level between POSTER (3^{rd} best from our previous study), POSTER++ (2^{nd} best from our previous study) and QUADCANN. The accuracy of QUADCANN outperformed that of both the POSTER and POSTER++ models for each emotion.

QUADCANN exhibited at least 3.8% and at maximum 21.2% greater accuracy for all emotions, except for contempt (2. 48%) and fear (2. 4%), as shown in Fig. 16. This is because of the lip features that we considered for classification. Eyes will show more signs of fear than lips can. Similarly, features with yaw angles are more indicative of contempt than lips [22]. As a result, we observe decreased accuracy for these two expressions.

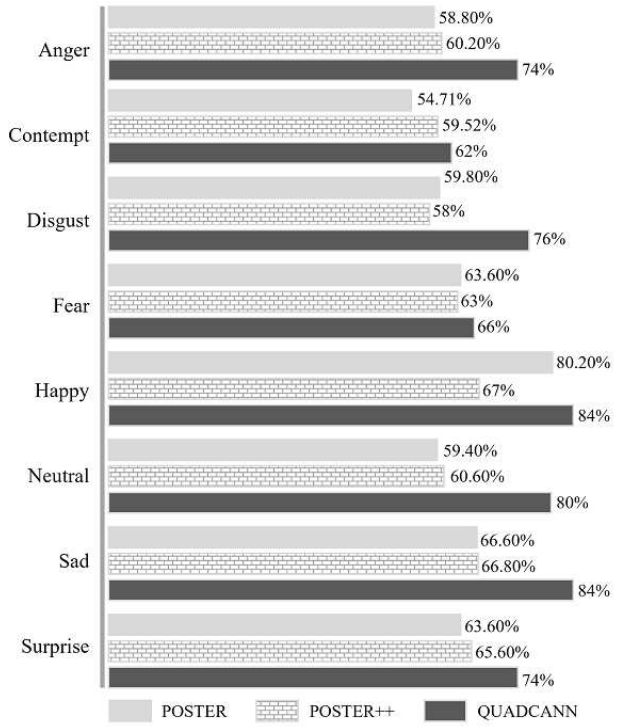
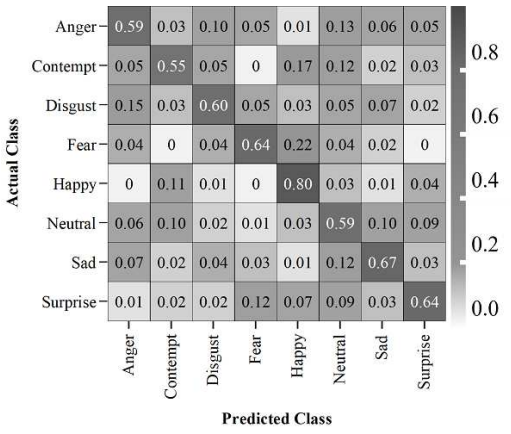


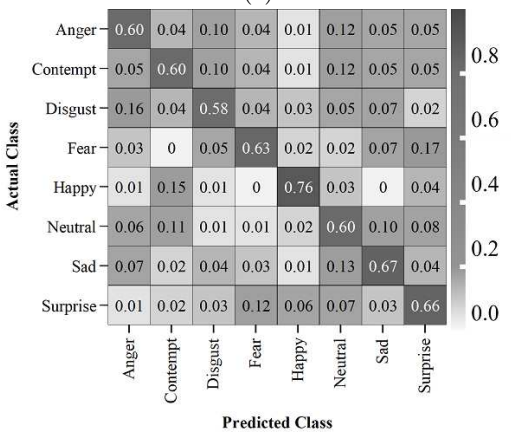
Fig. 16. Accuracy Comparison at the Emotional Level between POSTER, POSTER++, and QUADCANN

4.2.4 Confusion Matrix

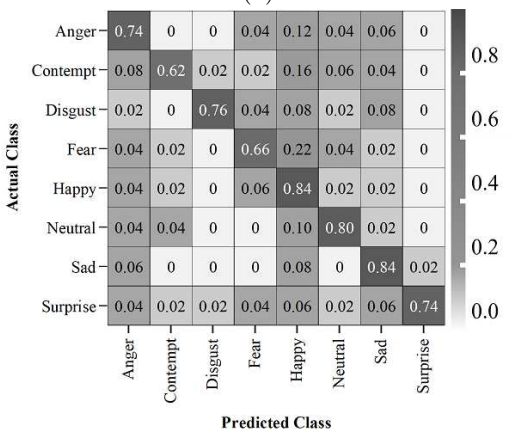
The normalized confusion matrices of the POSTER, POSTER++, and QUADCANN models are shown in Figs. 17a, 17b & 17c. The best classification accuracies from each model on the AffectNet-8 dataset are depicted in the figure. For the given predicted class, the zero value represents no misclassification. Compared to other models, QUADCANN has a greater number of zero values in both the upper and lower diagonal matrix, indicating a lower rate of misclassification. A lower degree of accuracy is the consequence of more classification errors, which are linked to higher values in the upper and lower diagonal matrix. The QUADCANN model's confusion matrix demonstrated that it had higher values only for contempt and fear, which were categorized as happy. We have provided an explanation for this observation in Section 4.2.3. The correct predictions are represented by diagonal values, and a higher value represents higher accuracy.



(a)



(b)

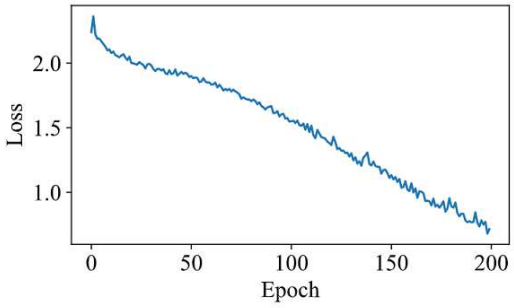


(c)

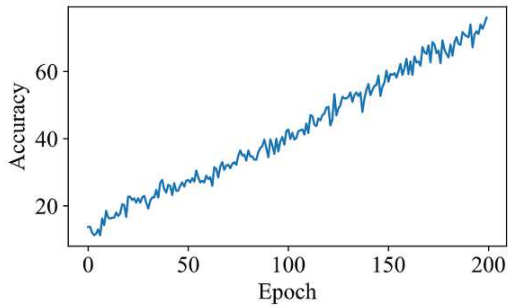
Fig. 17. Normalized Confusion Matrices of the SOTA models on AffectNet-8 Dataset (a) POSTER (b) POSTER++ (c) QUADCANN

4.2.5 Loss vs Accuracy

As we observed earlier in section 4.1, our QUADCANN model experienced a high initial average loss. However, as the model progresses through the training process, the average loss gradually decreases and the accuracy increases, as shown in Fig. 18a. The model achieved its maximum accuracy at 200 epochs, as shown in Fig. 18b.



(a)



(b)

Fig. 18. (a) Loss w.r.t. Epoch (b) Training Accuracy w.r.t. Epoch

5 Conclusion and Future Work

In summary, the QUADCANN model exhibits superior accuracy and performance compared to the CDC, QDC, and SOTA models in isolation. Although classical models are observed to deliver superior results for a limited number of features, quantum-classical hybrid models become more efficient as the number of features increases. QUADCANN demonstrated relatively higher accuracy compared to other state-of-the-art FEA models.

Further study is needed to determine the significance of RGB images for improving the classification accuracy of the models we highlighted. Similarly, additional improvements to the ANN models are required to increase efficiency, which will be a future focus of this work. We acknowledge that classification accuracy could be further enhanced by optimizing the model that incorporates quantum error as a key constraint. Encouraging results show promising future prospects for quantum computing and quantum image processing with remarkable advances in accuracy and performance. Our study affirms the possibility of improving classification accuracy using hybrid classifiers. Quantum algorithms considerably improve the performance of ML and DL models while handling complex feature extraction tasks. We hope that these findings will enable upcoming developments in quantum computing and machine learning.

Acknowledgements The authors sincerely thank TATA Consultancy Services for sponsoring this research.

The authors acknowledge the use of IBM quantum services [36] for this work. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM quantum team.

In this paper we used ibmq_osaka, which is one of the IBM Quantum Eagle R3 processors and ibmq_qasm_simulator.

References

1. Sarker, H, I.: Machine learning: Algorithms, real-world applications and research directions. SN Computer Science **2**, 1–21 (2021). DOI 10.1007/s42979-021-00592-x
2. Sonika, M., Hlavac, V., Boyle, R.: Image pre-processing: Image Processing, Analysis and Machine Vision. Springer New York, NY (1993). DOI 10.1007/978-1-4899-3216-7
3. Casat, G., Benenti, G.: Quantum computation and chaos: Encyclopedia of condensed matter physics. Elsevier Science **5**, 9–16 (2005)
4. Owayjan, M., Achkar, R., Iskandar, M.: Face detection with expression recognition using artificial neural networks. In: 2016 3rd Middle East Conference on Biomedical Engineering (MECBME), pp. 115–119 (2016). DOI 10.1109/MECBME.2016.7745421
5. Schuld, M., Fingerhuth, M., Petruccione, F.: Implementing a distance-based classifier with a quantum interference circuit. EPL (Europhysics Letters) **119**(6), 60002 (2017). DOI 10.1209/0295-5075/119/60002
6. Ruan, Y., Xue, X., Shen, Y.: Quantum image processing: Opportunities and challenges. Mathematical Problems in Engineering pp. 1–8 (2021). DOI 10.1155/2021/6671613
7. Dang, Y., Jiang, N., Hu, H., Ji, Z., Zhang, W.: Image classification based on quantum K-nearest-neighbor algorithm. Quantum Information Processing **17**, 1–18 (2018). DOI 10.1007/s11128-018-2004-9
8. Kukulski, R., Pawela, L., Puchala, Z.: On the probabilistic quantum error correction. IEEE Transactions on Information Theory **69**(7), 4620–4640 (2023). DOI 10.1109/TIT.2023.3254054
9. Djordjevic, I.: Chapter 7 - quantum error correction. In: I. Djordjevic (ed.) Quantum Information Processing and Quantum Error Correction, pp. 227–276. Academic Press, Oxford (2012). DOI 10.1016/B978-0-12-385491-9.00007-1

10. Razuri, J.G., Sundgren, D., Rahmani, R., Cardenas, A.M.: Automatic emotion recognition through facial expression analysis in merged images based on an artificial neural network. In: IEEE 2013 12th Mexican International Conference on Artificial Intelligence, pp. 85–96 (2013). DOI 10.1109/MICAI.2013.16
11. Nath, S.S., Mishra, G., Kar, J., Chakraborty, S., Dey, N.: A survey of image classification methods and techniques. In: International Conference on Control, Instrumentation, Communication and Computational Technologies, pp. 554–557 (2014). DOI 10.1109/ICCICCT.2014.6993023
12. Giuseppe, A., Fabrizio, F., Claudio, G., Claudio, V.: A comparison of face verification with facial landmarks and deep features. In: MMEDIA 2018, The Tenth International Conference on Advances in Multimedia At: Athens, Greece, pp. 1–6 (2018)
13. Fathallah, A., Abdi, L., Douik, A.: Facial expression recognition via deep learning. In: 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), pp. 745–750 (2017). DOI 10.1109/AICCSA.2017.124
14. Jain, D.K., Shamsolmoali, P., Sehdev, P.: Extended deep neural network for facial emotion recognition. Pattern Recognition Letters **120**, 69–74 (2019). DOI 10.1016/j.patrec.2019.01.008
15. Kamachi, M., Lyons, M., Gyoba, J.: The japanese female facial expression (jaffe) database. kasrl (1997)
16. Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, pp. 94–101 (2010). DOI 10.1109/CVPRW.2010.5543262
17. Yue, Z., Yanyan, F., Shangyou, Z., Bing, P.: Facial expression recognition based on convolutional neural network. In: 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS), pp. 410–413 (2019). DOI 10.1109/ICSESS47205.2019.9040730
18. Christopher, K., Cottrell, G.W.: Color-to-grayscale: Does the method matter in image recognition? PLoS ONE (2012). DOI 10.1371/journal.pone.0029740
19. Caraiman, S., Manta, V.: Image processing using quantum computing. In: 2012 16th International Conference on System Theory, Control and Computing (ICSTCC), pp. 1–6 (2012)
20. Mengoni, R., Incudini, M., Pierro, A.D.: Facial expression recognition on a quantum computer. Springer Nature Quantum Machine Intelligence **3**(8) (2021). DOI 10.1007/s42484-020-00035-5
21. Khurelsukh, B.: Hybrid quantum convolutional neural networks in tensorflow quantum. In: 2022 33rd Irish Signals and Systems Conference (ISSC), pp. 1–7 (2022). DOI 10.1109/ISSC55427.2022.9826145
22. Bhatt, A., Alam, T., Rane, K.P., Nandal, R., Malik, M., Neware, R., , Goel, S.: Quantum-inspired metaheuristic algorithms with deep learning for facial expression recognition under varying yaw angles. International Journal of Modern Physics **33**(04) (2022). DOI 10.1142/S0129183122500450
23. Majumdar, R., Sur-Kolay, S.: Special session: Quantum error correction in near term systems. In: 2020 IEEE 38th International Conference on Computer Design (ICCD), pp. 9–12 (2020). DOI 10.1109/ICCD50377.2020.00015
24. Swathi, M., Rudra, B.: Novel encoding method for quantum error correction. In: 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC), pp. 1001–1005 (2022). DOI 10.1109/CCWC54503.2022.9720880
25. Rengasamy, K., Joshi, P., Raveendra, V.V.S.: Hybrid facial expression analysis model using quantum distance-based classifier and classical support vector machine. In: 2023 11th International Symposium on Electronic Systems Devices and Computing (ESDC), vol. 1, pp. 1–6 (2023). DOI 10.1109/ESDC56251.2023.10149860
26. Zheng, C., Mendieta, M., Chen, C.: Poster: A pyramid cross-fusion transformer network for facial expression recognition. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3146–3155 (2023)
27. Mao, J.J., Xu, R., Yin, X., Chang, Y., Nie, B., Huang, A.: Poster v2: A simpler and stronger facial expression recognition network. ArXiv **abs/2301.12149** (2023)
28. Mollahosseini, A., Hasani, B., Mahoor, M.H.: Affectnet: A database for facial expression, valence, and arousal computing in the wild. IEEE Transactions on Affective Computing **10**(01), 18–31 (2019). DOI 10.1109/TAFFC.2017.2740923
29. Iliyasu, A.M., Dong, F., Hirota, K.: A flexible representation of quantum images for polynomial preparation, image compression and processing operations, quantum inf. Springer Quantum Information Processing **10**(1), 63–84 (2011). DOI 10.1007/s11128-010-0177-y
30. Raveendra, V.V.S., Subramanian, G.: A priori estimation of triangles in an arbitrary planar triangulation. Computers & Structures **52**(2), 215–218 (1994). DOI 10.1016/0045-7949(94)90274-7
31. Lezoray, O., Grady, L.: Image Processing and Analysis with Graphs Theory and Practice, 1 edn. CRC Press (2017)
32. Tkachenko, N.V.: Chapter 4 - Optical measurements, 1 edn. Elsevier Science (2006)
33. Agarap, A.F.: Deep learning using rectified linear units (ReLU). ResearchGate (2018)
34. Baldi, P.: Gradient descent learning algorithm overview: a general dynamical systems perspective. IEEE Transactions on Neural Networks **6**(1), 182–195 (1995). DOI 10.1109/72.363438
35. Zhao, K., Song, L., Zhang, Y., Pan, P., Xu, Y., Jin, R.: Ann softmax: Acceleration of extreme classification training. Proc. VLDB Endow. **15**(1), 1–10 (2021). DOI 10.14778/3485450.3485451
36. Qiskit Contributors: Qiskit: An open-source framework for quantum computing (2023). DOI 10.5281/zenodo.2573505