An Efficient Binary Harris Hawks Optimization based on Quantum SVM for Cancer Classification Tasks

Essam H. Houssein¹, Zainab Abohashima², Mohamed Elhoseny³, and Waleed M. Mohamed¹

¹ Faculty of Computers and Information, Minia University, Egypt.
² Faculty of Computer Science, Nahda University, Beni-Suef, Egypt.
³ Department of Computer Science, American University in the Emirates, Dubai, UAE.

Abstract. Cancer classification based on gene expression increases early diagnosis and recovery, but high-dimensional genes with a small number of samples are considered the major challenge in cancer classification tasks. Gene selection is a critical step in large-dimensional features to select the most informative and relevant subset of genes. The Gene selection problem utilizes metaheuristic techniques to extract the optimal subset of genes in microarray cancer datasets, such as particle swarm optimization (PSO), Harris hawks optimization (HHO), and grey wolf optimizer (GWO). Binary Harris hawk optimization (BHHO) technique, which mimics the behavior of the cooperative action of Harris hawks in nature, is recently proposed as one of these techniques. The biggest challenge of quantum hardware is the limited number of qubits, which restricts the use of quantum devices in real applications. The principal component analysis (PCA) is applied to reduce the selected genes to match the qubit numbers. In this work, a new hybrid quantum-kernel support vector machine (QKSVM) with BHHO called BHHO–PCA–QKSVM for cancer classification on a quantum simulator. This study aims to improve the microarray cancer prediction performance with the quantum kernel estimation based on the informative genes by BHHO. The quantum computer is used for estimation the kernel with the training data of the reduced genes and generation of the quantum kernel matrix. Besides, the classical computer is used for drawing the support vectors based on the quantum kernel matrix and applying the prediction stage. The colon and breast microarray datasets are used for evaluating the proposed approach performance with all genes and the selected genes. The proposed model enhances the overall performance of the two datasets. Also, the final results of the proposed model compared with different quantum feature maps (kernels) and classical radial basis function (RBF) kernel.

Keywords: Quantum-Kernel Support Vector Machines; Gene Expression; quantum kernels; Breast cancer; Colon cancer; Feature selection;
1 Introduction

Quantum computing recently emerged with machine learning (ML) to introduce a new field, called, quantum machine learning (QML) [1]. Quantum computing is applied to enhance classical ML. QML is introduced in many concepts, including QML, quantum-hybrid ML, and quantum-inspired ML [2]. Various ML algorithms have been proposed in quantum versions, such as quantum k-means clustering [3], quantum nearest-neighbor [4], and decision tree classifier [5]. The QSVM is introduced in many versions [6,7,8]. In this work, we focus on a quantum kernel-enhanced SVM. Kernel functions for ML are commonly used in pattern recognition and classification tasks. Kernel estimation has become expensive and difficult for classical devices with the data size increase. Quantum devices can easily estimate the kernel function with large-dimensional vectors; however, the current quantum hardware is limited by the available number of qubits.

With the high-dimensionality genes in microarray cancer data, the microarray datasets contain excessive and insignificant genes. So, the overall accuracy of the ML model can get largely poor. Therefore, feature selection is a vital step for selecting valuable genes and enhancing performance. The filter and wrapper methods are the main feature selection techniques for gene cancer data. The wrapper method is more significant in feature selection [9]. The gene selection process can be split as diversification of all available solutions. Each solution is evaluated by the ML algorithm (e.g., SVM) [10,11].

Several meta-heuristic algorithms based on feature selection have been proposed [12,13]. Harris hawk optimization (HHO) is recently presented to solve optimization problems. HHO has a proven outstanding performance in several real-time optimization compared to other metaheuristic algorithms [14] and is applied in many applications [15,16,17]. In addition, in large-scale wireless sensor networks, the optimal sink node placement based on HHO is proposed [18]. Meanwhile, ML in the healthcare domain has recently made progress. ML plays a critical role in solving most problems related to the healthcare field, including diagnosis and classification of medical data, extraction patterns in data, and detection of chronic diseases to avoid the risk of diseases. The most common chronic diseases are cancer, stroke, diabetes, and kidney and heart diseases. The selection of informative genes and the classification of high-dimensional features are some of the most important research topics. The number of genes in microarray data has a large dimensionality. Thus, the feature selection step is necessary with microarray datasets. This approach aims to

- Enhance the QKSVM performance with the quantum kernel estimation based on the all/selected genes by BHHO.
- Investigate the proposed approach performance with various quantum kernels, and
- Open the way for researchers to apply QML algorithms with real-world problems.

The key contributions of this work are as follows:
A hybrid QK SVM approach with PCA and BHHO-based feature selection for microarray cancer detection, called BHHO–PCA–QK SVM, is proposed. High-dimensional medical data are classified on a quantum simulator computer. The performance of microarray cancer classification with different quantum kernels is demonstrated.

The remainder of this paper is structured as follows: §2 introduces the hybrid approach BHHO–PCA–QK SVM in detail; §3 discusses the experimental results and provides a comparative performance with different quantum kernels; and finally, section §4 concludes this work.

2 The Proposed Approach

Due to the limited quantum hardware resources, the advantage of quantum SVM is not clear enough in real applications [19]. The proposed hybrid approach contains four steps: first, the BHHO selects informative genes; second, the PCA reduces the number of selected genes compatible with the number of qubits in the quantum device; third, the SMOTE method is used to handle the imbalanced classification problem; and fourth, the quantum kernel SVM is applied for the gene classification. Figure 1 shows the block diagram of the proposed approach.

![Block diagram of the proposed approach.](image)

2.1 Harris Hawks Optimization

The HHO is a new technique inspired by the chasing style proposed in 2019. The substantial idea of the HHO is based on mimicking the surprise pounce behavior. The surprise pounce is the cooperative action of Harris’ hawks in...
nature. The HHO consists of two phases: exploration and exploitation phases. In the exploration stage, the Harris hawks monitor and detect the victim (i.e., rabbit) using their sharp eyes. Harris hawks randomly settle in various spaces to attach to their prey. Mathematically, the exploration stage is defined as:

$$Z(t+1) = \begin{cases} 
Z_k(t) - r_1 (Z_k(t) - 2r_2 Z(t)) & e \geq 0.5 \\
(Z_r(t) - Z_m(t)) - r_3 (lb + r_4(ub - lb)) & e < 0.5 
\end{cases} \quad (1)$$

Where $Z$ is the Harris hawk location, and $t$ is the current iteration. The first law of the equation is to select the hawk location randomly. In the second law of the equation, $Z_r$ is the best solution (victim location); $Z_m$ is the mean hawk group location calculated from Eq(2), $r_1, r_2, r_3, r_4,$ and $e$ are the numbers at random $[0,1]$. $ub$ and $lb$ are the maximum and minimum boundaries, respectively.

$$Z_m(t) = \frac{1}{N} \sum_{n=1}^{N} Z_n(t) \quad (2)$$

Where $N$ is the Harris hawks number, and $X_n$ is the hawk number in the group. The change of hawk action relies on the escaping energy of the victim ($E$) to transform into the exploitation stage. This can be modeled as:

$$E = 2E_0 \left(1 - \frac{t}{T}\right), E_0 = 2r - 1 \quad (3)$$

Where $E_0$ is the random initial energy in $[-1, 1]$; $T$ is the total number of iterations; $r$ is number at random $[0,1]$, and $t$ is the current iteration. In the exploitation stage, the hawk location is varied based on the victim’s energy (escaping energy). This stage is performed following the four main steps of the smooth blockade, hard blockade, smooth blockade with a progressive quick pounce, and hard blockade with a progressive quick pounce.

**Binary Harris Hawks Optimization** Jingwei et al. [20] presented a new version of HHO for feature selection tasks, called binary HHO (BHHO). The BHHO algorithm is applied to the transfer function (e.g., S-shaped) to convert the HHO to binary feature selection technique. [21]. The Harris hawk location is updated in two points. The binary HHO updates the current location $Z_i^d(t)$ of the Harris hawk into a new location $\Delta Z_i^d(t + 1)$. The using V-/S-shaped function is used to apply the probability form of the HHO algorithm. The new hawk location using S-shaped functions is defined as follows:

$$Z_i^d(t + 1) = \begin{cases} 
1 & \text{if rand} \ (0,1) < T (\Delta Z_i^d(t + 1)) \\
0 & \text{otherwise} 
\end{cases} \quad (4)$$

Where

- $Z$ is the hawk location
- $i$ is the hawk number in the group
- $d$ is the dimension
- $t$ is the current iteration

The new location of the hawk by using V-shaped functions can be modeled as
follows:

\[ Z^d_i(t + 1) = \begin{cases} \neg X^d_i(t) & \text{if } \text{rand}(0, 1) < T(\Delta Z^d_i(t + 1)) \\ Z^d_i(t) & \text{otherwise} \end{cases} \]  

Where:

- \( Z \) is the hawk location
- \( i \) is the hawk number in the group
- \( t \) is the current iteration
- \( d \) is the dimension
- \( \neg X \) is the complement of \( X \)

### 2.2 Synthetic Minority Oversampling Technique

The synthetic minority oversampling technique (SMOTE) is a method to handle the imbalanced class problem in large dimensions’ data [22]. The key idea of SMOTE is to increase the number of observations in the minority class to be equal to the sample numbers in the majority class. In the colon dataset, the minority class (normal) has 22 samples. After using the SMOTE methods, the normal class has 31 samples. In the breast dataset, the minority class has 10 normal samples. After applying the SMOTE method, the normal class has 64 samples (as seen in Table 1).

### 2.3 Principal Component Analysis

Currently, one of the challenges of QML is the limited number of qubits, which restricts the use of quantum devices in applications with QML algorithms [2,23]. Therefore, using a quantum computer with complex large dimensionality data is still difficult. In this study, the PCA method is used to transform the selected gene number to match the available number of qubits, which still retains most of the gene information of the genes.

### 2.4 Quantum Kernel SVM

The SVM is a widespread classifier model broadly utilized in medical applications. The binary SVM works to discover the best hyperplane, which classifies the features into two categories +1, -1. The SVM model works to great the margin between the best hyperplane and the nearest data points, called the support vectors (as shown in Figure 2). In 2019, two quantum algorithms in supervised ML are proposed [24]: variational quantum classifier (VQC) and quantum kernel estimator (QKE). In VQC and QKE, the quantum state space is used as the feature space. The VQC is based on a quantum variation circuit described in [25,26] for data classification in a way like classical SVMs. The VQC consists of two stages (training and classification stage). In the training stage, the authors used four steps to compute the hyperplane between training data. The classification stage is used to classify new data with the correct label. The second technique, the main idea of QKE is to estimate a quantum kernel on quantum
device and using this kernel with classical SVMs. The key idea of QK SVM is to apply a quantum kernel that would be hard to simulate with classical devices. Quantum kernels map data from classical $\vec{x}$ into quantum state $|\Phi(\vec{x})\rangle$.

$$K(\vec{x}, \vec{z}) = |\langle \Phi(\vec{x}) | \Phi(\vec{z}) \rangle|^2$$  \hspace{2em} (6)

The steps of the hybrid QK SVM algorithms can be explained as follow:

- The quantum computer is used to estimate the kernel (applying the Equation 7) from the training data of the all/informative genes of the Colon/Breast dataset.
- The classical computer is used to obtain the support vectors based on the quantum kernel matrix.
- The prediction stage is performed with the classical device according to the obtained support vectors based-quantum kernel from the training data.

The three quantum feature maps (kernels) are used with the proposed approach: ZFeaturemap, ZZFeatureMap, and PauliFeatureMap (a combination of $Z$ and $YY$ matrices). The quantum feature map function can be generated from the following equation:

$$U_{\Phi(\vec{x})} = U_{\Phi(\vec{x})} H^\otimes n U_{\Phi(\vec{z})} H^\otimes n$$ \hspace{2em} (7)

Where $\mu_{\Phi(\vec{x})}$ is the Pauli expansion unitary, $H$ is the Hadamard gate, and $n$ is the qubit number.

$$\mu_{\Phi(\vec{x})} = \exp \left( i \sum_{S \subseteq [n]} \phi_S(\vec{x}) \prod_{i \in S} P_i \right)$$  \hspace{2em} (8)
3 Results and Discussion

This section presents the used microarray colon and breast datasets and analyzes the experimental results acquired through the proposed approach. The results are implemented on Google Colab by using the Qiskit framework. Qiskit is a cross-platform Python library for quantum computation. The quantum feature map is executed on a quantum computer simulator with 20 qubits, called the QASM simulator. The results are presented in two parts for clarity.

3.1 Microarray datasets

The proposed approach is evaluated using two microarray cancer datasets, namely Colon [27], and Breast [28] datasets. The total number of genes in the Colon dataset is 2000. There are 40 instances in the dataset that are assigned to the tumor and the other 22 instances are normal. Forty instances in the dataset are assigned to the tumor, and the other 22 instances are normal. The Breast dataset consists of 33,578 genes and 139 samples. Table 1 shows a description of the two used datasets and class distribution after using the SMOTE technique. Table 1 shows a description of the two used datasets and class distribution after using SMOTE technique.

| dataset | Genes | Samples | Class Distribution | Class Distribution after SMOTE |
|---------|-------|---------|--------------------|--------------------------------|
| Colon   | 2000  | 62      | 40 tumor 22 normal | 49 tumor 31 normal             |
| Breast  | 33578 | 139     | 129 primary 10 normal | 129 primary 64 normal          |

3.2 Evaluation Methods

The proposed approach performance is evaluated with five measures: accuracy, recall, precision, f1-measure, and receiver operating characteristic–area under curve (ROC–AUC) curve. The five measures can be calculated by a confusion matrix, which contains four elements (i.e., TP, TN, FP, and FN) defined as follows:

- TP: true positive; indicates the proposed model correctly predicts cancer cases and is classified as cancer cases
- TN: true negative; means the proposed model correctly predicts normal and is classified as normal
- FP: false positive; means the proposed model incorrectly predicts normal and is classified as normal
- FN: false negative; indicates the proposed approach incorrectly predicts cancer cases and is classified as a cancer case
Accuracy (Acc.) = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)

True Positive Rate / Recall (Re.) = \frac{TN}{TN + FP} \quad (10)

Precision (Pre.) = \frac{TP}{TP + FP} \quad (11)

F1-score = \frac{2 \times (Pre. \times Re.)}{Pre. + Re.} \quad (12)

False Positive Rate = \frac{FP}{FP + TN} \quad (13)

3.3 Results Analysis

Before the result analysis, Table 2 shows the parameters setting of the quantum kernels used in this work. The parameters of QK SVM are feature map, the number of shots is 100, the number of seed-simulator and seed-transpiler is 10,598. The first part of the results is applied with all genes and the selected genes by BHHO. The QK SVM algorithm is used with the ZZFeatureMap quantum kernel. First, the QK SVM algorithm with PCA is applied to two datasets with all genes. Finally, the selected genes by BHHO are used with the QK SVM-PCA. Table 3 presents the classification performance on the test data set with all genes and the selected genes by BHHO. The proposed approach based on selected genes by BHHO enhances the overall performance of the colon and Breast datasets. In two data sets, the proposed model achieved better performance in terms of accuracy, recall, f1-score compared with that using all genes. The proposed model obtains the highest accuracies of 87.5% and 95% for the colon and breast datasets, respectively. Moreover, the proposed approach achieves the highest AUC-ROC score in the two datasets as shown in Figure 3. The second part offers the results of the proposed approach with another two quantum kernels (ZFeatureMap, and PauliFeatureMap) and classical kernel (RBF). Table 4 presents the classification performance on the test data set with different functions. In the colon data, the classical SVM (RBF) outperformed the two kernel functions(ZFeatureMap and PauliFeatureMap). PauliFeatureMap is a combination of Z and YY Feature Maps. The SVM(RBF) obtains the highest accuracy of 93.7%, whereas the
Table 3
Classification performance based on all genes and the selected genes by BHBO

| Dataset | Approach                  | ACC.(%) | Pre.(%) | Re.(%) | F1-score(%) |
|---------|---------------------------|---------|---------|--------|-------------|
| Colon   | QSKVM+PCA                 | 83.3    | 100     | 71.4   | 83.3        |
|         | BHBO+PCA+QSKVM            | 87.5    | 100     | 87.5   | 88          |
| Breast  | QSKVM+PCA                 | 90      | 83.3    | 100    | 90          |
|         | BHBO+PCA+QSKVM            | 95      | 90.9    | 100    | 95.2        |

QSKVM (PauliFeatureMap) obtains the lowest accuracy of 75%. Also, the highest AUC is achieved by classical SVM as offered in Figure 4. In Breast data, the QSKVM (ZFeatureMap) achieves the highest accuracy with 95%. By contrast, The QSKVM (PauliFeatureMap) obtains the lowest accuracy with 85%. Besides, the highest accuracy is obtained by QSKVM (ZFeatureMap) as shown in Figure 5. The quantum advantage is not noticed owing to the limited number of qubits and the current quantum hardware. The quantum advantage display in complex problems, in which classical computers solve Rebentrost et al. proposed a quantum version of the SVM for big data classification. Currently, this version does not run on quantum devices.

Table 4
Classification performance based on different quantum kernels and classical RBF kernel.

| Dataset | Algorithm | Kernel      | ACC.(%) | Pre.(%) | Re.(%) | F1-score(%) |
|---------|-----------|-------------|---------|---------|--------|-------------|
| Colon   | QSKVM     | ZFeatureMap | 81.2    | 100     | 75     | 85.7        |
|         |           | PauliFeatureMap | 75     | 83.3    | 83.3   | 83.3        |
|         | SVM       | RPF         | 93.7    | 100     | 91.6   | 95.6        |
| Breast  | QSKVM     | ZFeatureMap | 95      | 90      | 100    | 93.2        |
|         |           | PauliFeatureMap | 85     | 76.9    | 100    | 86.9        |
|         | SVM       | RPF         | 92.5    | 86.9    | 100    | 93          |

4 Conclusion and Future Work

Binary Harris hawks optimization is integrated with quantum kernel SVM for selecting the genes and classification of microarray datasets. The proposed approach aims to improve the microarray cancer classification performance based on the quantum kernel estimation with the informative genes. The two microarray datasets are applied for evaluating the proposed approach with all genes and the selected genes. The principal component analysis is applied to reduce the informative genes to match the qubit numbers of the quantum simulator. Various
Fig. 3: The AUC-ROC curve of the proposed approach and QKSVM+PCA for (a) Colon and (b) Breast datasets.

(a)  

(b)  

Fig. 4: The AUC-ROC curve of the proposed approach with different quantum kernel and classical kernel (RBF) (a) Colon and (b) Breast datasets.

(a)  

(b)
quantum feature maps are used with the QK SVM to convert data into quantum feature space. The proposed approach archives the higher accuracy with the selected genes 87.5% and 95% for colon and breast cancer, respectively. Also, the performance proposed approach is compared with different quantum kernels and classical kernel RBF. Due to the limited quantum hardware resources, the advantage of quantum SVM is not clear enough in real applications [19,29]. In future work, more quantum kernels can be applied with real quantum devices. The variational quantum classifier can also be used with different datasets.

References

1. J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” Nature, vol. 549, no. 7671, pp. 195–202, 2017.
2. Z. Abohashima, M. Elhosen, E. H. Houssein, and W. M. Mohamed, “Classification with quantum machine learning: A survey,” arXiv preprint arXiv:2006.12270, 2020.
3. S. Lloyd, M. Mohseni, and P. Rebentrost, “Quantum algorithms for supervised and unsupervised machine learning,” arXiv preprint arXiv:1307.0411, 2013.
4. N. Wiebe, A. Kapoor, and K. M. Svore, “Quantum nearest-neighbor algorithms for machine learning,” Quantum information and computation, vol. 15, 2018.
5. S. Lu and S. L. Braunstein, “Quantum decision tree classifier,” Quantum information processing, vol. 13, no. 3, pp. 757–770, 2014.
6. Z. Li, X. Liu, N. Xu, and J. Du, “Experimental realization of a quantum support vector machine,” Physical review letters, vol. 114, no. 14, p. 140504, 2015.
7. P. Rebentrost, M. Mohseni, and S. Lloyd, “Quantum support vector machine for big data classification,” Physical review letters, vol. 113, no. 13, p. 130503, 2014.
8. D. Willsch, M. Willsch, H. De Raedt, and K. Michielsen, “Support vector machines on the d-wave quantum annealer,” Computer physics communications, vol. 248, p. 107006, 2020.
9. K. Hussain, N. Neggaz, W. Zhu, and E. H. Houssein, “An efficient hybrid sine-cosine harris hawks optimization for low and high-dimensional feature selection,” Expert Systems with Applications, vol. 176, p. 114778, 2021.

10. M. Al-Rajab, J. Lu, and Q. Xu, “Examining applying high performance genetic data feature selection and classification algorithms for colon cancer diagnosis,” Computer methods and programs in biomedicine, vol. 146, pp. 11–24, 2017.

11. E. H. Houssein, D. S. Abdelminiaam, H. N. Hassan, M. M. Al-Sayed, and E. Nabil, “A hybrid barnacles mating optimizer algorithm with support vector machines for gene selection of microarray cancer classification,” IEEE Access, vol. 9, pp. 64895–64905, 2021.

12. A. G. Hussien, A. E. Hassanien, E. H. Houssein, S. Bhattacharyya, and M. Amin, “S-shaped binary whale optimization algorithm for feature selection,” in Recent trends in signal and image processing. Springer, 2019, pp. 79–87.

13. N. Neggaz, E. H. Houssein, and K. Hussain, “An efficient henry gas solubility optimization for feature selection,” Expert Systems with Applications, vol. 152, p. 113364, 2020.

14. A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, “Harris hawks optimization: Algorithm and applications,” Future generation computer systems, vol. 97, pp. 849–872, 2019.

15. E. H. Houssein, M. E. Hosney, M. Elhoseny, D. Oliva, W. M. Mohamed, and M. Hassaballah, “Hybrid harris hawks optimization with cuckoo search for drug design and discovery in chemoinformatics,” Scientific Reports, vol. 10, no. 1, pp. 1–22, 2020.

16. F. A. Hashim, E. H. Houssein, K. Hussain, M. S. Mabrouk, and W. Al-Atabany, “A modified henry gas solubility optimization for solving motif discovery problem,” Neural Computing and Applications, vol. 32, no. 14, pp. 10759–10771, 2020.

17. E. H. Houssein, M. E. Hosney, D. Oliva, W. M. Mohamed, and M. Hassaballah, “A novel hybrid harris hawks optimization and support vector machines for drug design and discovery,” Computers & Chemical Engineering, vol. 133, p. 106656, 2020.

18. E. H. Houssein, M. R. Saad, K. Hussain, W. Zhu, H. Shaban, and M. Hassaballah, “Optimal sink node placement in large scale wireless sensor networks based on harris’ hawk optimization algorithm,” IEEE Access, vol. 8, pp. 19381–19397, 2020.

19. F. Phillipson, R. S. Wezeman, and I. Chiscop, “Three quantum machine learning approaches or mobile user indoor-outdoor detection,” in 3rd International Conference on Machine Learning for Networking (MLN’2020)(online), 2020.

20. J. Too, A. R. Abdullah, and N. Mohd Saad, “A new quadratic binary harris hawk optimization for feature selection,” Electronics, vol. 8, no. 10, p. 1130, 2019.

21. S. Saremi, S. Mirjalili, and A. Lewis, “How important is a transfer function in discrete heuristic algorithms,” Neural Computing and Applications, vol. 26, no. 3, pp. 625–640, 2015.

22. A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla, “Smote for learning from imbalanced data: progress and challenges, marking the 15-year anniversary,” Journal of artificial intelligence research, vol. 61, pp. 863–905, 2018.

23. S. Y.-C. Chen, C.-M. Huang, C.-W. Hsing, and Y.-J. Kao, “An end-to-end trainable hybrid classical-quantum classifier,” arXiv preprint arXiv:2102.02416, 2021.

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

25. K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, “Quantum circuit learning,” Physical Review A, vol. 98, no. 3, p. 032309, 2018.
26. E. Farhi and H. Neven, “Classification with quantum neural networks on near term processors,” *arXiv preprint arXiv:1802.06002*, 2018.
27. A. B. Olshen and A. N. Jain, “Deriving quantitative conclusions from microarray expression data,” *Bioinformatics*, vol. 18, no. 7, pp. 961–970, 2002.
28. B. C. Feltes, E. B. Chandelier, B. I. Grisci, and M. Dorn, “Cumida: an extensively curated microarray database for benchmarking and testing of machine learning approaches in cancer research,” *Journal of Computational Biology*, vol. 26, no. 4, pp. 376–386, 2019.
29. E. H. Houssein, Z. Abohashima, M. Elhoseny, and W. M. Mohamed, “Hybrid quantum convolutional neural networks model for covid-19 prediction using chest x-ray images,” *arXiv preprint arXiv:2102.06585*, 2021.