Received: November 7, 2024. Revised: November 24, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Hybrid Quantum-Deep Learning Approach: Optimizing Land Cover Classification with GMM Outlier and Fusion Key Feature Selection

T. Sutojo[1] **De Rosal Ignatius Moses Setiadi**[1]* **Supriadi Rustad**[1] **Muhamad Akrom**[1]
Wise Herowati[1]

Research Center for Quantum Computing and Materials Informatics, Faculty of Computer Science, Dian Nuswantoro University, Semarang 50131, Indonesia
* Corresponding author's Email: moses@dsn.dinus.ac.id

Abstract: This study proposes a hybrid Quantum-Deep Learning framework to enhance the efficiency and accuracy of land cover classification. The approach integrates a Quantum Gaussian Mixture Model (QGMM) for outlier detection, a Bidirectional Gated Recurrent Unit (BiGRU) for feature extraction, and a combination of Random Forest (RF) and XGBoost (XGB) for feature selection. By leveraging quantum principles in the Expectation-Maximization (EM) algorithm, the QGMM significantly optimizes parameter estimation, enabling robust outlier detection that supports better model generalization. Experimental results on the UCI Forest Cover dataset demonstrate near-perfect classification performance, achieving an accuracy of 99.9% and an Area under the curve (AUC) of 1.0, highlighting the model's capacity to handle high-dimensional, imbalanced data effectively. This framework provides a promising solution for complex environmental datasets, paving the way for future research into integrating quantum techniques for broader ecological and geospatial applications.

Keywords: Hybrid quantum model, Land cover classification, Quantum clustering, Quantum expectation-maximization, Quantum Gaussian mixture model.

Table 1. Notation List

Notation	Definition
x	Data point in the dataset
N	Total number of data points in the dataset
K	Number of Gaussian components in the GMM
π_k	Weight or prior probability of the k-th Gaussian component
μ_k	Mean of the k-th Gaussian component
Σ_k	Covariance matrix of the k-th Gaussian component
γ_{ik}	Posterior probability that data point x_i belongs to the k-th Gaussian component
$\mathcal{N}(x\|\mu_k, \Sigma_k)$	Gaussian distribution function
H	Hadamard gate for creating superposition in the quantum circuit
θ	Rotation angle for quantum gates based on Gaussian parameters
RZ	Quantum rotation gates around the z-axis
RX	Quantum rotation gates around the x-axis
RY	Quantum rotation gates around the y-axis
$CNOT$	Controlled-NOT gate for entanglement in quantum computing
ψ_k	Quantum state probability amplitude for component k
cost	Cost function used for optimizing parameters in Quantum EM
$FI_{RF}(j)$	Feature importance score for feature j calculated by Random Forest
$\Delta I_t(j)$	Impurity reduction for feature j in tree t

$FI_{XGB}(j)$	Feature importance score for feature j calculated by XGBoost
T	Total number of trees in the Random Forest or XGBoost ensemble
gain	Reduction in impurity or loss after splitting on a feature in XGBoost
$\text{loss}_{\text{before}}, \text{loss}_{\text{after}}$	Loss values before and after a split in XGBoost

1. Introduction

Accurate land cover classification is essential for various environmental applications, such as natural resource management, land use planning, and climate change mitigation. Accurate land cover data helps researchers and policymakers make better decisions regarding ecosystem conservation and environmental management [1–3]. However, reliable classification requires complex data processing and efficient approaches to handle large and heterogeneous datasets.

One of the biggest challenges in land cover classification is extracting features that can capture hidden patterns and relationships between attributes. Deep learning models like recurrent neural networks have shown great potential in solving such tasks [4]. However, these approaches have limitations in handling data that require more complex bidirectional processing. A newer alternative, the Bidirectional Gated Recurrent Unit (BiGRU), offers the advantage of processing bidirectional sequential data, allowing the model to combine information from the past and future simultaneously, enhancing feature representation [5]. Thus, BiGRU provides significant advantages in capturing richer patterns compared to unidirectional approaches such as LSTM and traditional GRU [6].

Although these approaches are effective, there is still a need to improve efficiency and enrich the extracted information, which can be achieved with quantum computing-based approaches. The integration of quantum computing in feature extraction allows for improvements in terms of processing and deeper feature capture, thanks to its ability to leverage the principles of quantum superposition and interference [7–9].

On the other hand, outlier detection in data is also significantly critical since outliers can interfere with model training and lead to biased or inaccurate predictions [10, 11]. The application of outlier detection has been proven effective in improving model performance in various fields, including agriculture [12], data security [13], business [14], and healthcare [15]. Integrating outlier detection into the data analysis pipeline can help improve data quality before the model is trained, resulting in more stable and reliable classification results.

Classical methods, such as Isolation Forest, One-Class SVM, and Local Outlier Factor (LOF), have been widely used in various studies to detect outliers [16]. However, these methods have scale and computational efficiency limitations, especially for large and complex datasets. Meanwhile, the Gaussian Mixture Model (GMM) has an Expectation-Maximization (EM) algorithm [17, 18]. EM is an algorithm widely used in statistical model parameter estimation, especially when latent variables exist. The iterative nature and modular structure of EM make it relatively easy to modify and implement in quantum computing [19, 20]. Thus, Quantum GMM (QGMM) can offer a solution to process parameter optimization more efficiently than classical methods [21]. This allows faster parameter space exploration and reduces the risk of getting stuck in local optima [22].

The increase in the number of features resulting from the feature engineering process, both deep learning-based feature extraction techniques and quantum computing, as well as outlier detection, often causes increased computational complexity, especially in high-dimensional datasets [7, 13, 14]. Whereas not all features are necessarily useful and make the dataset more informative. Therefore, feature selection is an important step in ensuring that only the most informative features are retained, reducing computational complexity while maintaining or even improving model performance [6, 23, 24]. Various feature selection methods such as Recursive Feature Elimination (RFE), Information Gain(IG), and Chi-square have been applied to address this issue [25, 26]. However, these methods may miss essential features that contribute to model accuracy. The combination of Random Forest (RF) and XGBoost (XGB) has emerged as a practical approach in feature selection, combining the stability and in-depth evaluation provided by RF with the high sensitivity of XGB in identifying significant features [27, 28]. This combination enables more informative feature selection, reducing complexity without sacrificing model performance.

In recent years, quantum computing approaches have attracted attention as a potential solution to complex optimization and computational challenges. Unlike classical computing, quantum computing offers advantages in solving optimization problems more efficiently [21, 29]. This study provides new opportunities to improve model performance in land cover classification and maximize computational efficiency by integrating quantum computing methods with deep learning and classical machine learning techniques.

This paper proposes a Hybrid Quantum-Deep Learning framework that combines Quantum GMM for outlier detection, feature extraction using BiGRU, and feature selection based on a combination of Random Forest and XGBoost. This framework aims to optimize the land cover classification process by improving the overall accuracy and computational efficiency. The proposed method integrates quantum computing, BiGRU, and RF-XGB to optimize land cover classification. Quantum computing accelerates parameter estimation in GMM, BiGRU extracts temporal patterns, and RF-XGB reduces computational complexity, forming a robust framework for high-dimensional and imbalanced datasets.

The rest of this paper is organized as follows: Section 2 reviews related work to position this study within the current research landscape. Section 3 details the proposed hybrid framework, including preprocessing, feature extraction, outlier detection, and feature selection. Section 4 presents the results and analysis, highlighting the performance and robustness of the proposed method. Finally, Section 5 concludes the study and outlines future research directions.

2. Related works

Land cover classification has become an important research subject due to its relevance in natural resource management and ecosystem balance. Previous studies have explored the use of classical machine learning algorithms, such as K-Nearest Neighbors (KNN), Random Forest (RF), Gradient boosting, and Support Vector Machine (SVM), for land cover dataset classification from the UCI Machine Learning Repository [30–32]. These studies highlight the effectiveness of supervised learning methods in processing structured datasets with diverse features, including cartographic variables such as elevation, slope, soil type, and distance to hydrology.

Studies have shown that RF and KNN often excel in terms of accuracy. For example, RF consistently outperformed other algorithms in classifying forest cover types due to its ability to handle high-dimensional data and reduce overfitting through ensemble learning [33, 34]. Using non-parametric methods such as KNN has proven effective, with significant accuracy results of up to 97.09%, significantly better than the baseline values reported in the UCI Repository [33]. However, these methods face challenges when applied to large datasets, particularly regarding computational efficiency. KNN, for instance, requires extensive distance computations, making it less practical for real-time applications.

Additionally, ensemble-based studies, such as the EMLARDE method, have highlighted the advantages of ensemble models in improving multiclass classification accuracy [35]. By employing a decorrelation mechanism among decision trees, EMLARDE achieves enhanced performance for multiclass datasets. Despite their effectiveness, ensemble methods are computationally intensive, especially when dealing with high-dimensional data or datasets with significant imbalances in class distributions. These limitations restrict their scalability for large-scale or real-time classification problems [31, 34, 35].

Traditional optimization methods often struggle with complex classification problems due to their reliance on iterative and deterministic approaches, which may fail to find global optima in high-dimensional spaces. Recently, quantum computing methods have been introduced to overcome these challenges. Techniques such as Annealing Lévy Quantum Inspired Particle Swarm Optimization (ALQPSO) have demonstrated advantages in parameter space exploration, offering higher convergence speeds and better optimization in diverse applications [36]. However, their application to land cover classification remains limited.

Although previous studies have investigated feature selection methods, outlier detection, and machine learning algorithms for land cover classification, significant gaps remain. The integration of quantum computing methods for optimization, Quantum GMM-based outlier detection, and deep learning models such as BiGRU have not been thoroughly explored. Existing methods often fail to address the combined challenges of noisy data and imbalanced class distributions. This study fills this gap by proposing a Hybrid Quantum-Deep Learning framework that integrates QGMM, BiGRU, and RF-XGB for feature selection to improve land cover classification accuracy and computational efficiency.

3. Proposed method

The proposed method in this study combines quantum approaches, deep learning, and classical methods to improve the accuracy and efficiency of land cover classification. This approach involves several stages, from data preprocessing feature extraction using deep learning architectures such as BiGRU, outlier detection based on GMM and

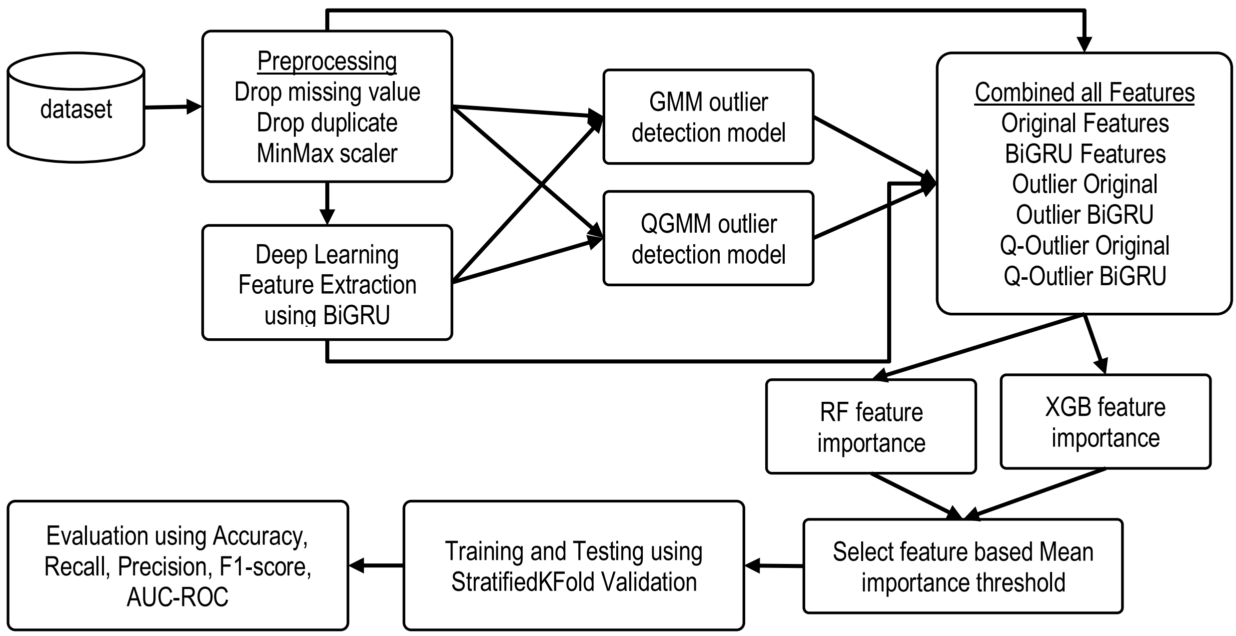

Figure. 1 Overview proposed method

QGMM, and feature selection with Random Forest and XGBoost.

3.1 Preprocessing

The preprocessing process ensures that the data is in optimal condition before being used in the model. The first step involves removing missing and duplicate data. Discrete data labels are converted into a numeric format using LabelEncoder so that the machine learning model can process them. Next, numeric features are scaled using MinMaxScaler to normalize the data between 0 and 1, ensuring that all features have the same scale and thus avoiding bias in the algorithm.

3.2 Deep Learning Future Extraction

Feature extraction uses the Bidirectional Gated Recurrent Unit (BiGRU) architecture. BiGRU is a variant of Recurrent Neural Network (RNN) that processes data in two directions, forward and backward, so it can simultaneously capture contextual information from the past and future. The BiGRU model used in this study is implemented in the Sequential framework with the following parameters:

1. Input Layer to receive input data. This section requires dimension expansion to be compatible with the BiGRU layer, namely data transformation from $(n_{\text{samples}}, n_{\text{features}})$ to $(n_{\text{samples}}, n_{\text{features}}, 1)$
2. BiGRU Layer with 16 units in each direction, equipped with return_sequences=False, ensures that only the layer's last output is returned.
3. The dense layer is a fully connected layer with eight neurons and a ReLU activation function, which helps reduce non-linearity and processes the BiGRU output into a denser feature vector.

The BiGRU integration steps in feature extraction ensure that temporal and complex patterns in the dataset are well captured so that the model can produce more informative feature representations for the next stage.

3.3 Quantum-Classic Outlier Detection

The classical Gaussian Mixture Model (GMM) and QGMM are used at this stage. GMM is a probabilistic model that models the data distribution as a mixture of several Gaussian distributions. The probability that data x is generated by a Gaussian mixture distribution with K k components is given by Eq. (1).

$$p(x|\theta) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k) \tag{1}$$

Where π_k is the weight or proportion of the k^{th} Gaussian component; μ_k is the mean of the k^{th} component; Σ_k is the k^{th} component covariance matrix; $\mathcal{N}(x|\mu_k, \Sigma_k)$ is a Gaussian distribution function described in Eq. (2).

$$\mathcal{N}(x|\mu_k, \Sigma_k) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) \tag{2}$$

Expectation-Maximization (EM) is used to estimate the parameters θ of GMM in two iterative stages, namely Expectation Step (E-Step) and Maximization Step (M-Step). E-Step functions to calculate the latent probability distribution for data x_i, while M-Step is responsible for updating the parameters based on the latent probability distribution. E-step and M-step can be calculated by Eq. (3) and (4), respectively.

$$\gamma_{ik} = \frac{\pi_k \mathcal{N}(x_i|\mu_k, \Sigma_k)}{\sum_{j=1}^{K}(x_i|\mu_j, \Sigma_j)} \quad (3)$$

$$\mu_k = \frac{\sum_{i=1}^{N}\gamma_{ik}x_i}{\sum_{i=1}^{N}\gamma_{ik}}, \quad \Sigma_k = \frac{\sum_{i=1}^{N}\gamma_{ik}(x_i-\mu_k)(x_i-\mu_k)^T}{\sum_{i=1}^{N}\gamma_{ik}} \quad (4)$$

To modify the EM algorithm into a Quantum version, several quantum circuits are designed to perform E-step and M-step calculations in a quantum state, allowing simultaneous processing of multiple parameters. In this study, a quantum circuit is designed with 4 qubit inputs processed through several layers of processing, consisting of the following quantum gates:

1. A Hadamard (H) gate is applied to each qubit at the beginning of the circuit to place the qubit in a superposition state. This allows each qubit to have an equal chance of being in the states $|0\rangle$ and $|1\rangle$. The Hadamard gate is expressed as in Eq. (5).
2.

$$H = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (5)$$

3. After the initial superposition, the RY and RZ rotation gates are applied to adjust the quantum state based on the Gaussian distribution parameters. Each qubit is rotated about the Y and Z axes using angles corresponding to the model parameters, such as the Gaussian mean and covariance. The RY and RZ rotations are performed by Eq. (6) and (7), respectively. The RY gate controls the amplitude of the qubit state, while the RZ gate controls the phase.

$$RY(\theta) = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{bmatrix} \quad (6)$$

$$RZ(\theta) = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix} \quad (7)$$

4. The CNOT gate is applied between adjacent qubits in the third layer to form entanglement. Entanglement allows strong interactions between qubits, which is useful in capturing correlations between variables in multivariable data. The CNOT gate acts as a control gate, changing the state of the target qubit only if the control qubit is in the $|1\rangle$ state. The CNOT matrix is given as in Eq. (8).

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (8)$$

5. The last layer is a measurement after rotation with the RX gate. This gate performs an additional rotation on the qubit, setting the final orientation before the measurement. The probability measurement calculates the probability distribution used in the parameter updates at the E-step and M-step. The matrix for rotation around the X-axis is given in Eq. (9).

$$RY(\theta) = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) & -i\sin\left(\frac{\theta}{2}\right) \\ -i\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{bmatrix} \quad (9)$$

For a clearer visualization of the quantum gates, see Figure 2. Passing through these quantum stages makes the Quantum E-step calculated in a quantum state that allows simultaneous evaluation for all Gaussian components (see Eq. (10)). The parameter update in the M-step is done by utilizing the measurement results of the quantum circuit to update the mean and covariance, as in Eq. (11). Finally, to perform quantum EM optimization, a cost function is added to maximize the parameter distribution shown in Eq. (12).

$$\gamma_{ik}^{\text{quantum}} = \frac{|\psi_k\rangle|^2}{\sum_{j=1}^{K}|\psi_k\rangle|^2} \quad (10)$$

$$\mu_k^{\text{quantum}} = \frac{\sum_{i=1}^{N}\gamma_{ik}^{\text{quantum}}x_i}{\sum_{i=1}^{N}\gamma_{ik}^{\text{quantum}}} \quad (11)$$

$$\text{cost} = -\sum_{\text{measured}}|\psi_k\rangle^2 \quad (12)$$

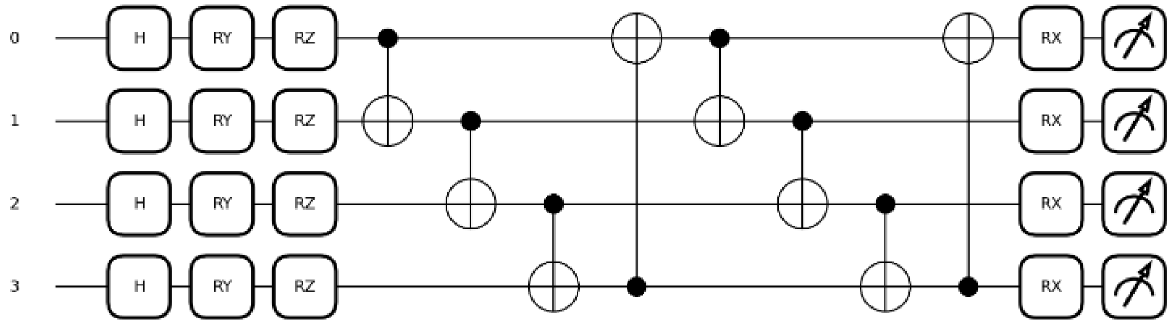

Figure. 2. Quantum circuit visualization

Once the cost values are calculated in the Quantum EM optimization step, these values are used to update the parameters in the quantum circuit through the optimization step. In the given code, the optimization is performed using AdamOptimizer, which updates the parameters based on the resulting cost values. The cost values are not directly fed into the circuit but are used as a guide to direct the parameter updates to achieve the optimal probability distribution. This approach speeds up EM iterations by reducing computational complexity, allowing efficient and precise parameter processing for large-dimensional datasets.

3.4 Fusion feature selection

Feature selection is done by combining the evaluation results from RF and XGB. RF calculates feature importance with Eq. (13). While XGB evaluates features based on the gain generated by each division with Eq. (14). The gain value is then accumulated for each occurrence of feature j in all trees in the XGB ensemble, thus providing the final feature importance value $FI_{XGB}(j)$. The higher the gain value generated by the feature, the more critical the model considers the feature.

$$FI_{RF}(j) = \sum_{t=1}^{T} \frac{\Delta I_t(j)}{T} \tag{13}$$

$$\text{gain} = \text{loss}_{\text{before}} - \text{loss}_{\text{after}} \tag{14}$$

Where $\Delta I_t(j)$ is the impurity reduction in tree t for feature j.

Finally, the mean importance is calculated by combining the scores from both models in Eq. (15).

$$\text{mean importance}_j = \frac{FI_{RF}(j) + FI_{XGB}(j)}{2} \tag{15}$$

Features with mean importance above a threshold are selected to reduce computational complexity while maintaining model accuracy.

3.5 Training and Testing

In this study, a 5-fold StratifiedKFold was used, where the dataset was divided into five subsets, each of which had a similar class distribution. In each iteration, four subsets were used as training data, and one subset was used as testing data. This process was repeated five times so that each subset became testing data once and training data four times. The results of these five folds were then averaged to produce more stable evaluation metrics and reduce variability in model performance.

3.6 Evaluation

Several evaluation metrics are used to assess the model's overall performance, namely Accuracy, Precision, Recall, F1-Score, and Area Under Curve (AUC). Accuracy provides a general view of the model's performance but can be less informative in imbalanced data. Precision is important to measure the reliability of the model's positive predictions. Recall is relevant in cases where identifying all positive classes is more important than avoiding false positive predictions. The F1 score evaluates the model's overall performance in handling the positive class without being too influenced by the class distribution. AUC is the area under the Receiver-operating characteristic (ROC) curve, whose value ranges between 0 and 1. An AUC value close to 1 indicates that the model can distinguish between positive and negative classes. To understand more details about how to calculate accuracy, precision, recall, f1-score, and AUC, you can see Eq. (16)-(20).

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

$$\text{precision} = \frac{TP}{TP + FP} \tag{17}$$

$$\text{recall} = \frac{TP}{TP + FN} \tag{18}$$

$$\text{f1} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (19)$$

$$\text{AUC} = \int_0^1 \text{ROC curve} \quad (20)$$

4. Results and analysis

4.1 Experimental setup

Experiments were conducted on Google Colab using Python and core libraries such as TensorFlow and Keras for deep learning model development. Scikit-Learn was used for preprocessing, cross-validation, and metric evaluation, while XGBoost and Random Forest were applied for feature selection. The quantum Gaussian Mixture Model (QGMM) was developed using the PennyLane quantum simulator. Data visualization and analysis results were performed using Matplotlib and Seaborn. The dataset used was taken from[37], which has seven classes with a total of 581012 records. The class distribution is very imbalanced, as can be clearly seen in Figure 3. This dataset has 54 features, so after deep learning and outlier processing both classically and quantumly, it has 112 features, consisting of 54 original features, 54 BiGRU features, original classic outliers, original quantum outliers, BiGRU classic outliers, BiGRU quantum outliers.

This dataset has also been proven to have quite dominant outliers, as evidenced by the feature

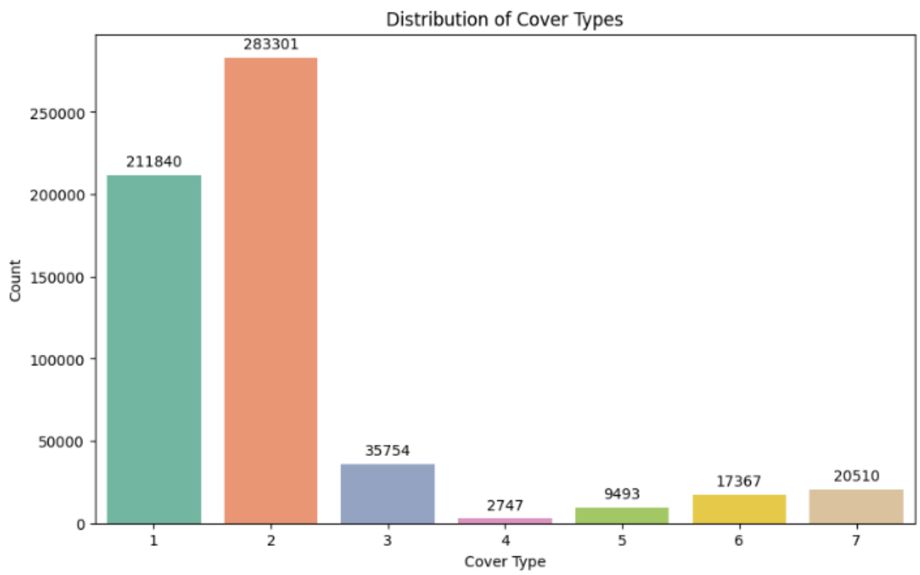

Figure. 3 Cover type class distribution

Table 2. Selected Features

No	Feature	No	Feature
1	BiGRU_QuantumGMM_Outlier_Label	15	Soil_Type39
2	BiGRU_GMM_Outlier_Label	16	Wilderness_Area4
3	QuantumGMM_Outlier_Label_Original	17	Wilderness_Area1
4	GMM_Outlier_Label_Original	18	Soil_Type10
5	BiGRU_Feature_0	19	Aspect
6	Soil_Type4	20	Soil_Type11
7	BiGRU_Feature_2	21	Soil_Type38
8	BiGRU_Feature_6	22	Hillshade_9am
9	BiGRU_Feature_3	23	Horizontal_Distance_To_Fire_Points
10	Elevation	24	Soil_Type32
11	BiGRU_Feature_7	25	Horizontal_Distance_To_Roadways
12	Soil_Type40	26	Hillshade_3pm
13	BiGRU_Feature_5	27	Soil_Type23
14	Soil_Type2	28	Horizontal_Distance_To_Hydrology

Table 3. Classification Results

Metrics	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Acc	1.0	1.0	0.99999	1.0	1.0
Prec	1.0	1.0	0.99999	1.0	1.0
Recall	1.0	1.0	0.99999	1.0	1.0
F1	1.0	1.0	0.99999	1.0	1.0
AUC-ROC	1.0	1.0	1.0	1.0	1.0

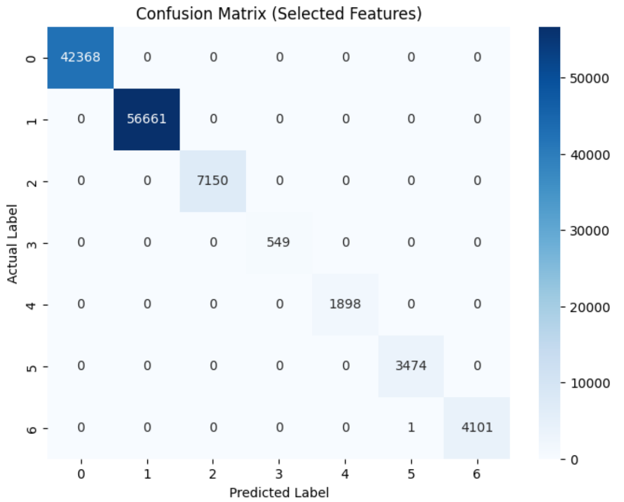

Figure. 4 Confusion matrix in fold 3

selection process using RF and XGB, where all four outlier features are included in the important feature category. The selected features are features that have a threshold mean importance > 0.001. The total selected features are 28 features, presented in Table 2.

4.2 Results

After the feature selection process, the classification results with 5-fold cross-validation are presented in Table 3. Meanwhile, the sample confusion matrix is presented in Figure 4. Based on the results presented in Table 3, this method achieves almost perfect performance on every major evaluation metric, such as accuracy, precision, recall, and F1-score, with an average of 999998 for all folds in cross-validation. This shows that the proposed method is very efficient in handling complex and high-dimensional land cover classification, thanks to the integration of quantum processing, deep learning, and ensemble-based feature selection.

The achievement of perfect AUC-ROC values across all folds demonstrates the model's ability to distinguish classes consistently and accurately without experiencing performance degradation. This reliability is mainly supported by the utilization of

Table 4. Ablation Study

Method	Acc	Prec	Recall	F1	AUC-ROC
Without feature selection	0.957	0.957	0.957	0.957	0.998
Without QGMM	0.947	0.947	0.947	0.947	0.997
Proposed	0.999	0.999	0.9999	0.999	1.0

Quantum EM on GMM, which allows the model to accelerate parameter convergence without getting stuck in local optima and precise feature selection to reduce data dimensions without sacrificing important information.

The confusion matrix visualization in Fold 3 shows that the model experiences almost no misclassification. This strengthens the conclusion that the model can classify with great precision, even on class-imbalanced data. This is supported by feature analysis, which shows that all outlier features are included in the important feature category.

4.3 Comparison and analysis

In this section, we perform some analysis related to the performance of the proposed method with the ablation studies presented in Table 4. Table 3 shows the ablation study results evaluating the proposed model's performance by removing the main components, namely feature selection and QGMM. The results without feature selection show a decrease, indicating that RF-XGB feature selection helps reduce less relevant features and improves model generalization by focusing on the most informative attributes.

Furthermore, removing QGMM results in a greater decrease in performance, with accuracy, precision, recall, F1 score of 0.947, and AUC-ROC of 0.997. QGMM plays an essential role in detecting outliers and optimizing model parameters. This method helps the model handle deviant observations and capture complex patterns in the data, which supports classification accuracy. This decrease in performance reinforces the importance of QGMM in improving the model's ability to handle complex datasets.

The proposed model, which combines QGMM and feature selection, achieves the highest performance with perfect scores on all metrics (accuracy, precision, recall, F1, and AUC-ROC of 1.0). This shows that an integrative approach that utilizes the advantages of QGMM, BiGRU for feature extraction, and RF-XGB for feature selection has significantly improved the classification performance.

Received: November 7, 2024. Revised: November 24, 2024.

Table 5. Comparison with related works

Ref	Acc	Prec	Recall	F1	AUC-ROC
Ref [30]	0.966	-	-	-	-
Ref [31]	0.965	0.965	0.965	0.965	-
Ref [32]	0.93	0.93	0.92	0.92	-
Ref [34]	0.946	-	-	-	-
Ref[35]	≈0.91	-	-	-	-
Proposed	0.999	0.999	0.9999	0.999	1.0

Overall, both QGMM and feature selection contribute to the robustness and accuracy of the model, which are essential to achieving optimal results in land cover classification. Finally, a comparison was also made with several other studies, as shown in Table 5.

5. Conclusions

This study demonstrates the effectiveness of integrating quantum-enhanced methods with classical machine learning and deep learning techniques for land cover classification. The proposed approach achieved improved accuracy and computational efficiency by employing a QGMM for outlier detection, BiGRU for feature extraction, and combining feature selection with Random Forest and XGBoost. The experimental results show a near-perfect classification performance with an average accuracy of 99.99%, precision of 99.99%, recall of 99.99%, F1-score of 99.99%, and AUC-ROC of 1.0 across five cross-validation folds. These metrics highlight the model's capability to effectively handle high-dimensional and imbalanced data.

The quantum enhancements, particularly through applying the Quantum EM (QEM) algorithm, proved valuable in optimizing the parameter estimation process, contributing to more accurate and reliable classification results. For instance, the ablation study reveals that removing QGMM significantly drops accuracy to 94.7%, underlining the importance of quantum-based outlier detection in this framework. Moreover, feature selection using RF-XGB reduces computational complexity by selecting only 28 features from the original 112 without compromising performance, as demonstrated by the retained high evaluation metrics.

This fusion of quantum and classical methods highlights the potential of quantum computing in handling complex data distributions and large-scale environmental datasets, demonstrating the approach's applicability in real-world scenarios. Future work could further explore extending this framework to other types of datasets, such as those in urban planning or ecological monitoring, to validate its versatility. Additionally, enhancing the quantum circuit design by optimizing gate operations or increasing qubit numbers could improve computational efficiency and result precision, especially as quantum hardware continues to advance.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, TS and DRIMS; methodology, TS and DRIMS; software, MA; validation, SR, MA, and WH; formal analysis, SR; investigation, DRIMS; resources, TS; data curation, TS; writing—original draft preparation, TS; writing—review and editing, All; visualization, MA; supervision, SR; project administration, WH; funding acquisition, All.

References

[1] M. C. Hansen *et al.*, "High-Resolution Global Maps of 21st-Century Forest Cover Change", *Science (80-.).*, Vol. 342, No. 6160, pp. 850–853, 2013, doi: 10.1126/science.1244693.

[2] Food and Agriculture Organization of the United Nations, *Global Forest Resources Assessment 2020*, 2020. [Online]. Available: https://www.fao.org/interactive/forest-resources-assessment/2020/en/

[3] I. Daniel, L. Akinyemi, and O. Udekwu, "Identifying Landslide Hotspots Using Unsupervised Clustering: A Case Study", *J. Futur. Artif. Intell. Technol.*, Vol. 1, No. 3, pp. 249–268, 2024, doi: 10.62411/faith.3048-3719-37.

[4] W. Long, Z. Lu, and L. Cui, "Deep learning-based feature engineering for stock price movement prediction", *Knowledge-Based Syst.*, Vol. 164, pp. 163–173, 2019, doi: 10.1016/j.knosys.2018.10.034.

[5] Z. Hu, G. Liu, Y. Li, and S. Zhuang, "SAGB: self-attention with gate and BiGRU network for intrusion detection", *Complex Intell. Syst.*, Vol. 10, No. 6, pp. 8467–8479, 2024, doi: 10.1007/s40747-024-01577-y.

[6] D. R. I. M. Setiadi, S. Widiono, A. N. Safriandono, and S. Budi, "Phishing Website Detection Using Bidirectional Gated Recurrent Unit Model and Feature Selection", *J. Futur. Artif. Intell. Technol.*, Vol. 2, No. 1, pp. 75–83, 2024, doi: 10.62411/faith.2024-15.

[7] D. R. I. M. Setiadi, A. Susanto, K. Nugroho, A.

R. Muslikh, A. A. Ojugo, and H. Gan, "Rice Yield Forecasting Using Hybrid Quantum Deep Learning Model", *Computers*, Vol. 13, No. 8, p. 191, 2024, doi: 10.3390/computers13080191.

[8] M. Akrom *et al.*, "Quantum machine learning for corrosion resistance in stainless steel", *Mater. Today Quantum*, Vol. 3, No. August, p. 100013, 2024, doi: 10.1016/j.mtquan.2024.100013.

[9] A. N. Safriandono, D. R. I. M. Setiadi, A. Dahlan, F. Z. Rahmanti, I. S. Wibisono, and A. A. Ojugo, "Analyzing Quantum Feature Engineering and Balancing Strategies Effect on Liver Disease Classification", *J. Futur. Artif. Intell. Technol.*, Vol. 1, No. 1, pp. 51–63, 2024, doi: 10.62411/faith.2024-12.

[10] X. Cai, B. Aydin, A. Ji, and R. Angryk, "A Framework for Local Outlier Detection from Spatio-Temporal Trajectory Datasets", In: *Proc. of 2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 5682–5689, 2021, doi: 10.1109/ICPR48806.2021.9412274.

[11] D. Chakraborty, V. Narayanan, and A. Ghosh, "Integration of deep feature extraction and ensemble learning for outlier detection", *Pattern Recognit.*, Vol. 89, pp. 161–171, 2019, doi: 10.1016/j.patcog.2019.01.002.

[12] A. R. Muslikh, P. N. Andono, A. Marjuni, and H. A. Santoso, "Ensemble IDO Method for Outlier Detection and N2O Emission Prediction in Agriculture", *Int. J. Adv. Comput. Sci. Appl.*, Vol. 15, No. 7, 2024, doi: 10.14569/IJACSA.2024.0150737.

[13] Z. S. Dhahir, "A Hybrid Approach for Efficient DDoS Detection in Network Traffic Using CBLOF-Based Feature Engineering and XGBoost", *J. Futur. Artif. Intell. Technol.*, Vol. 1, No. 2, pp. 174–190, 2024, doi: 10.62411/faith.2024-33.

[14] D. R. I. M. Setiadi, A. R. Muslikh, S. W. Iriananda, W. Warto, J. Gondohanindijo, and A. A. Ojugo, "Outlier Detection Using Gaussian Mixture Model Clustering to Optimize XGBoost for Credit Approval Prediction", *J. Comput. Theor. Appl.*, Vol. 2, No. 2, pp. 244–255, 2024, doi: 10.62411/jcta.11638.

[15] "Enhancing Prediction Accuracy in an Imbalanced Dataset of Dengue Infection Cases Using a Two-layer Ensemble Outlier Detection and Feature Selection", *Int. J. Intell. Eng. Syst.*, Vol. 17, No. 2, pp. 544–560, 2024, doi: 10.22266/ijies2024.0430.44.

[16] E. F. Agyemang, "Anomaly detection using unsupervised machine learning algorithms: A simulation study", *Sci. African*, Vol. 26, p. e02386, 2024, doi: 10.1016/j.sciaf.2024.e02386.

[17] J. You, Z. Li, and J. Du, "A new iterative initialization of EM algorithm for Gaussian mixture models", *PLoS One*, Vol. 18, No. 4, p. e0284114, 2023, doi: 10.1371/journal.pone.0284114.

[18] M. R. Gupta, "Theory and Use of the EM Algorithm", *Found. Trends® Signal Process*, Vol. 4, No. 3, pp. 223–296, 2010, doi: 10.1561/2000000034.

[19] I. Kerenidis, A. Luongo, and A. Prakash, "Quantum Expectation-Maximization for Gaussian Mixture Models", *arXiv*. Aug. 19, 2019. [Online]. Available: http://arxiv.org/abs/1908.06657.

[20] H. Miyahara, K. Aihara, and W. Lechner, "Quantum expectation-maximization algorithm", *Phys. Rev. A*, Vol. 101, No. 1, p. 012326, 2020, doi: 10.1103/PhysRevA.101.012326.

[21] J. Preskill, "Quantum Computing in the NISQ era and beyond", *Quantum*, Vol. 2, p. 79, 2018, doi: 10.22331/q-2018-08-06-79.

[22] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum Support Vector Machine for Big Data Classification", *Phys. Rev. Lett.*, Vol. 113, No. 13, p. 130503, 2014, doi: 10.1103/PhysRevLett.113.130503.

[23] "A Hybrid Deep Learning Technique for Feature Selection and Classification of Chronic Kidney Disease", *Int. J. Intell. Eng. Syst.*, Vol. 16, No. 6, pp. 638–649, 2023, doi: 10.22266/ijies2023.1231.53.

[24] J. A. Ingio, A. S. Nsang, and A. Iorliam, "Optimizing Rice Production Forecasting Through Integrating Multiple Linear Regression with Recursive Feature Elimination", *J. Futur. Artif. Intell. Technol.*, Vol. 1, No. 2, pp. 96–108, 2024, doi: 10.62411/faith.2024-17.

[25] M. I. Akazue, I. A. Debekeme, A. E. Edje, C. Asuai, and U. J. Osame, "UNMASKING FRAUDSTERS: Ensemble Features Selection to Enhance Random Forest Fraud Detection", *J. Comput. Theor. Appl.*, Vol. 1, No. 2, pp. 201–211, 2023, doi: 10.33633/jcta.v1i2.9462.

[26] F. Omoruwou, A. A. Ojugo, and S. E. Ilodigwe, "Strategic Feature Selection for Enhanced Scorch Prediction in Flexible Polyurethane Form Manufacturing", *J. Comput. Theor. Appl.*, Vol. 1, No. 3, pp. 346–357, 2024, doi: 10.62411/jcta.9539.

[27] T. Chen and C. Guestrin, "XGBoost", In: *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.

[28] L. Breiman, “Random Forests”, *Mach. Learn.*, Vol. 45, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[29] A. W. Harrow, A. Hassidim, and S. Lloyd, “Quantum Algorithm for Linear Systems of Equations”, *Phys. Rev. Lett.*, Vol. 103, No. 15, p. 150502, 2009, doi: 10.1103/PhysRevLett.103.150502.

[30] D. Patel, A. Gupta, A. Pal, H. Parmar, and M. Gidwani, “Forest Cover Type Prediction”, In: *Proc. of 2023 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, pp. 1–5, 2023, doi: 10.1109/ICACTA58201.2023.10393537.

[31] S. J. Suhail and J. Refonaa, “Forest Cover Type Classification Using Machine Learning Algorithms”, In: *Proc. of 2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, pp. 1–5, 2023, doi: 10.1109/ICDCECE57866.2023.10151111.

[32] M. Mohammed Al Sameer, T. Prasanth, and R. Anuradha, “Rapid Forest Cover Detection Using Ensemble Learning”, In: *Proc. of International Virtual Conference on Industry 4.0*, pp. 181–190, 2021, doi: 10.1007/978-981-16-1244-2_15.

[33] P. Maurya and A. Kumar, “Performance Assessment of K-Nearest Neighbor Algorithm for Classification of Forest Cover Type”, In: *Proc. of Third International Conference on Sustainable Computing*, pp. 43–51, 2022, doi: 10.1007/978-981-16-4538-9_5.

[34] A. Kumar and N. Sinha, “Classification of Forest Cover Type Using Random Forests Algorithm”, *Advances in Data and Information Sciences*, Vol. 94, pp. 395–402, 2020, doi: 10.1007/978-981-15-0694-9_37.

[35] T. Guhan and N. Revathy, “EMLARDE tree: ensemble machine learning based random de-correlated extra decision tree for the forest cover type prediction”, *Signal, Image Video Process.*, Vol. 18, No. 12, pp. 8525–8536, 2024, doi: 10.1007/s11760-024-03470-0.

[36] P. Agarwal, A. Sahoo, and D. Garg, “An Improved Quantum Inspired Particle Swarm Optimization for Forest Cover Prediction”, *Ann. Data Sci.*, Vol. 11, No. 6, pp. 2217–2233, 2024, doi: 10.1007/s40745-023-00509-w.

[37] UCI Machine Learning, “Forest Cover Type Dataset,” *Kaggle*, 2016. https://www.kaggle.com/datasets/uciml/forest-cover-type-dataset (accessed Nov. 01, 2024).