To appear in:

International Journal of Mathematical Modelling & Computations

Online ISSN: 2228-6233 Print ISSN: 2228-6225

Received: 30- January-2026

Revised: 19- May -2026

Accepted: 24- May-2026

DOI: https://doi.org/10.57647/ijm2c.2027.1701.06

ORIGINAL RESEARCH

Hybrid Optimization of K-Means Clustering for BIG Data Using Deep Reinforcement Learning and Quantum Whale Optimization Algorithm

Zaid Hikmat Kareem Albayati[1], Mahdi Mosleh[2*], Yahya Mahdi Hadi Al-Mayali[3], Mohammad Mosleh[4]

1- Department of Computer Engineering, Isf.C., Islamic Azad University, Isfahan, Iran.

2*- Department of Computer Engineering, Isf.C., Islamic Azad University, Isfahan, Iran.

3- Department of Computer Techniques Engineering, College of Technical Engineering University of Alkafeel, Najaf, Iraq
4- Department of Computer Engineering, Dez. C., Islamic Azad University, Dezful, Iran

zaid.hikmatalbayati@iau.ac.ir , [2*]Mahdi_mosleh@iau.ac.ir, [3] yahya.almayali@alkafeel.edu.iq,

[4]Mohammad.mosleh@iau.ac.ir

Correspondence: Mahdi_mosleh@iau.ac.ir

Abstract

In this paper, a novel hybrid clustering framework, Deep Reinforcement Learning–Quantum Whale Optimization–KMeans (DRL–QWOA–KMeans), is proposed to overcome major limitations of conventional K-Means, including initialization sensitivity, premature convergence, and reduced clustering accuracy on complex data distributions. In the proposed model, Deep Reinforcement Learning (DRL) is employed to dynamically balance the exploration–exploitation process within the Quantum Whale Optimization Algorithm (QWOA), which adaptively optimizes the initial centroids of K-Means to achieve global convergence. Extensive experiments conducted on benchmark UCI datasets demonstrate the robustness and efficiency of the proposed method compared with classical meta-heuristic-based clustering algorithms such as GA, PSO, DE, WOA, and QWOA. Quantitative evaluation using multiple performance metrics—WCSS, AAcc, ASen, ASpe, and FscoreM—indicates a significant improvement in intra-cluster compactness and inter-cluster separability. Moreover, visual analysis confirms that DRL–QWOA–KMeans yields more stable and interpretable clusters compared to standard K-Means. These findings highlight the competitive potential of DRL-guided quantum optimization for complex clustering problems and provide a foundation for future investigations on larger-scale or streaming data environments.

Keywords : Clustering; Quantum-inspired Whale Optimization Algorithm (QWOA); Deep Reinforcement Learning (DRL); K-Means; Metaheuristics; Adaptive Optimization; Within-Cluster Sum of Squares (WCSS); Data Mining

Introduction

Data clustering plays a crucial role in unsupervised machine learning and knowledge discovery, enabling the identification of inherent structures within datasets [1]. Over the past five decades, the K-Means algorithm has been a cornerstone of clustering techniques due to its simplicity, scalability, and applicability across various domains, including image processing, bioinformatics, and text analysis [1,2,3]. However, traditional K-Means is plagued by limitations such as sensitivity to initial centroid placement, convergence to local optima, and inefficiency in handling outliers or imbalanced data [6,7]. These challenges are exacerbated in large-scale clustering problems, where high dimensionality and data volume demand more robust optimization strategies [4,5]. To address these shortcomings, researchers have increasingly turned to metaheuristic algorithms, which provide global search capabilities to enhance clustering performance [8]. Recent reviews highlight the efficacy of metaheuristics in optimizing K-Means, demonstrating improvements in convergence speed and solution quality [8,9]. For example, hybrid models integrating evolutionary and swarm intelligence techniques, such as Fruit-Fly Optimization with K-Means for text document clustering [11], Multi-Verse Optimizer with K-Means [12], and Gray Wolf Optimization hybridized with Grasshopper Optimization for feature selection and clustering [13], have shown promising results in complex applications.

Among swarm-based metaheuristics, the Whale Optimization Algorithm (WOA) has gained attention for its bio-inspired mechanisms, with enhancements applied to industrial parameter identification and simulation [10]. Quantum-inspired extensions of WOA (QWOA) further amplify its potential by incorporating quantum principles like superposition and entanglement, leading to superior exploration in feature selection and global optimization tasks. Similarly, the Capuchin Search Algorithm (CSA), a novel metaheuristic mimicking capuchin monkey behavior, has been successfully hybridized for nonlinear modeling [14,15], medical image analysis [16], intrusion detection with deep learning [17], wind power prediction [18], feature selection in COVID-19 diagnostics [19], plant image thresholding [20], and economic dispatch problems [21]. While these metaheuristics excel in static optimization, integrating adaptive learning mechanisms can further enhance their dynamism. Reinforcement learning (RL) has been combined with WOA variants to enable self-learning and parameter adaptation in scheduling and path planning. However, the fusion of quantum-enhanced WOA (QWOA) with Deep Reinforcement Learning (DRL) remains unexplored, particularly for large-scale clustering problems**.**

This paper introduces a novel hybrid optimization framework for K-Means clustering, leveraging QWOA for efficient global search in centroid initialization and DRL for adaptive parameter tuning. This approach mitigates local optima traps and scales effectively to large-scale clustering problems, outperforming traditional and existing hybrid methods. The main contributions of this study are summarized as follows. First, a novel hybrid clustering framework called DRL–QWOA–KMeans is proposed, which integrates Deep Reinforcement Learning with the Quantum-inspired Whale Optimization Algorithm to improve centroid initialization in K-Means clustering. Second, unlike conventional metaheuristic-based clustering approaches with fixed control parameters, the proposed framework employs a DRL agent to dynamically adjust QWOA parameters during the optimization process, enabling an adaptive balance between exploration and exploitation. Third, the parameter tuning of QWOA is formulated as a sequential decision-making problem where the DRL agent learns optimal parameter adjustment strategies based on clustering performance feedback. Finally, extensive experiments conducted on ten.

The main contributions of this study are summarized as follows:

1. We propose a novel hybrid framework, DRL–QWOA–KMeans, which integrates Deep Reinforcement Learning (DRL) with the Quantum-inspired Whale Optimization Algorithm (QWOA) to enable adaptive centroid initialization for K-Means clustering.
2. Unlike conventional metaheuristic-based clustering approaches that rely on fixed or manually tuned parameters, the proposed framework employs a Deep Reinforcement Learning agent to dynamically adjust the key control parameters of QWOA throughout the optimization process. This mechanism allows for adaptive balancing between exploration and exploitation according to the current search state.
3. We formulate the parameter adaptation of QWOA as a sequential decision-

making problem within a Markov Decision Process (MDP). The DRL agent learns effective parameter adjustment policies based on real-time feedback from clustering performance, thereby going beyond static heuristic tuning.

4. Extensive experiments conducted on ten benchmark datasets from the UCI repository demonstrate that the proposed method consistently outperforms several state-of-the-art optimization-based K-Means variants, including GA-KMeans, PSO-KMeans, DE-KMeans, WOA-KMeans, and QWOA-KMeans, across multiple evaluation metrics such as WCSS, AAcc, ASen, ASpe, and FscoreM. Statistical significance of the improvements was validated using the Wilcoxon signed-rank test.

To the best of our knowledge, this work is the **first study** that integrates Deep Reinforcement Learning for adaptive parameter control within a Quantum-inspired Whale Optimization Algorithm specifically for data clustering.

The remainder of the paper is organized as follows: Section 2 reviews related work, Section 3 details the proposed methodology, Section 4 presents experimental results, and Section 5 concludes with future directions.

2. Related Work

Recent years have witnessed extensive research on improving clustering performance through optimization techniques, particularly to overcome the inherent limitations of classical K-Means such as sensitivity to initialization, convergence to local optima, and poor scalability in complex and large-scale datasets. Researchers have explored a wide range of metaheuristic, swarm-intelligence, quantum-inspired, and learning-based approaches. Although these studies have shown improvements in clustering quality, most existing methods rely heavily on static parameter settings and lack adaptive learning mechanisms.

Several metaheuristic algorithms have been hybridized with K-Means. Karaboga and Ozturk [31] proposed an Artificial Bee Colony (ABC)-based clustering algorithm that outperformed classical K-Means. Shelokar et al. [32], Chen and Ye [33], Senthilnath et al. [34], and Hatamlou et al. [35–37] applied Ant Colony Optimization, Particle Swarm Optimization, Firefly Algorithm, Gravitational Search Algorithm, Black Hole, and Big Bang–Big Crunch algorithms, respectively. While these methods enhanced global search capability and clustering accuracy, they generally suffer from premature convergence, high sensitivity to control parameters, and limited adaptability across different datasets [38,39]. Quantum-inspired approaches have been introduced to further improve exploration capability. Boushaki et al. [40] proposed a quantum chaotic cuckoo search algorithm for data clustering. Although such methods increase population diversity, they still lack dynamic parameter adaptation during the optimization process.

In the domain of image processing and medical diagnosis, fuzzy clustering and deep learning methods have been widely used. Kumar et al. [24], Zhao et al. [25], and Dolz et al. [26] presented advanced fuzzy C-means and deep convolutional approaches for segmentation tasks. Gupta et al. [45] developed evolutionary multiobjective fuzzy clustering. However, these techniques are often supervised, computationally expensive, and less suitable for general unsupervised large-scale clustering problems. Other works focused on specific applications such as Braik et al. [22,23] with Chameleon Swarm Algorithm, and various hybrids involving Capuchin Search Algorithm [14–21]. Table 1 provides a comparative summary of representative studies from different categories, highlighting their key strengths and main limitations.

As summarized in Table 1, although significant efforts have been devoted to enhancing K-Means clustering, critical limitations persist. Most approaches, including those based on GA, PSO, DE, WOA, GSA, Firefly, and quantum-inspired methods [31–47], rely on static parameter settings and do not incorporate adaptive mechanisms to dynamically balance exploration and exploitation during the search. Deep learning and fuzzy-based methods, while effective in domain-specific tasks, are typically supervised and computationally intensive. To the best of our knowledge, the integration of Deep Reinforcement Learning for real-time adaptive parameter control within the Quantum-inspired Whale Optimization Algorithm (QWOA) for K-Means clustering has not been previously explored. The proposed **DRL–QWOA–KMeans** framework addresses this research gap by introducing a learning-based adaptive mechanism that dynamically regulates QWOA parameters based on the current optimization state and clustering performance feedback. This novel integration aims to achieve more robust exploration–exploitation balancing and superior clustering results across heterogeneous datasets. Although several studies have attempted to improve K-Means clustering using metaheuristic algorithms such as Genetic Algorithm (GA), Particle Swarm

Optimization (PSO), Differential Evolution (DE), Whale Optimization Algorithm (WOA), and Quantum-inspired Whale Optimization Algorithm (QWOA), most existing approaches rely on fixed or manually tuned control parameters during the optimization process. Such static parameter settings may lead to inefficient exploration–exploitation trade-offs and increase the likelihood of premature convergence.

In contrast, the proposed DRL–QWOA–KMeans framework introduces a learning-based adaptive parameter control mechanism. By employing a Deep Reinforcement Learning agent to dynamically regulate QWOA parameters based on the optimization state, the proposed method enables data-driven exploration–exploitation balancing and enhances the robustness of centroid search across different datasets. To the best of our knowledge, the integration of DRL-based adaptive control with QWOA for clustering optimization has not been previously explored. Recent research has also explored adaptive and learning-based optimization strategies, including reinforcement learning-assisted metaheuristics and automated machine learning (AutoML) techniques for clustering parameter optimization. These approaches aim to dynamically adjust algorithm parameters during the search process in order to improve convergence behavior and solution quality. Motivated by this research direction, the present study proposes a DRL-assisted parameter adaptation mechanism for the QWOA-based clustering framework.

Table 1. Comparative analysis of existing K-Means clustering enhancement methods

Category	Representative Studies	Key Strengths	Main Limitations
Classical K-Means and Extensions	[27]–[30]	Simple and fast implementation	Sensitive to initialization, local optima
Nature-Inspired Metaheuristics	[31]–[39], [41]–[44], [47]	Strong global search capability	Static parameters, limited scalability
Quantum-Inspired Algorithms	[22], [23], [40]	High exploration and population diversity	No adaptive learning control
Deep Learning & Fuzzy Methods	[24]–[26], [45]	High accuracy in domain-specific tasks	High computational cost, supervised
Hybrid Optimization-Based Methods	[38], [39]	Improved convergence and clustering quality	Increased complexity, static control

Although various hybrid metaheuristic approaches have been proposed to enhance K-Means clustering, most existing methods rely on static parameter settings and lack mechanisms for dynamic adaptation during the search process. Furthermore, while some studies have explored reinforcement learning-assisted metaheuristics in other domains, the integration of Deep Reinforcement Learning for real-time adaptive control of Quantum-inspired Whale Optimization parameters in the context of clustering remains unexplored. The proposed DRL–QWOA–KMeans framework addresses this research gap by introducing a learning-based adaptive mechanism that dynamically regulates the exploration–exploitation trade-off, leading to more robust and effective centroid optimization.

3. Proposed Method

3.1 Proposed Hybrid DRL–QWOA–KMeans Method

The main novelty of the proposed framework lies in incorporating a Deep Reinforcement Learning (DRL) agent into the Quantum-inspired Whale Optimization Algorithm (QWOA) to dynamically regulate its control parameters during the optimization process. Unlike conventional hybrid clustering approaches where the parameters of metaheuristic algorithms remain fixed or manually tuned, the proposed DRL–QWOA–KMeans framework enables adaptive parameter adjustment based on the current optimization state. This learning-based mechanism improves the balance between exploration and exploitation and enhances the global search capability of the algorithm.

In this section, a novel hybrid clustering framework is proposed to enhance the performance of K-Means for large-scale and high-dimensional data. The proposed method integrates Deep Reinforcement Learning (DRL) with a Quantum-inspired Whale Optimization Algorithm (QWOA) to address the main limitations of conventional K-Means, including sensitivity to initial centroid selection, static parameter settings, and susceptibility to local optima. Within the proposed framework, QWOA is utilized to perform an effective global search for optimal cluster centroids, while DRL dynamically regulates the key optimization parameters in order to maintain an appropriate balance between exploration and exploitation throughout the optimization process.

By combining the global exploration capability of quantum-inspired swarm intelligence with the adaptive decision-making mechanism of DRL, the proposed approach aims to achieve faster convergence, improved clustering accuracy, and enhanced robustness in data environments. The overall workflow of the proposed hybrid framework is illustrated in Figure 1 and is described in detail in the following subsections.

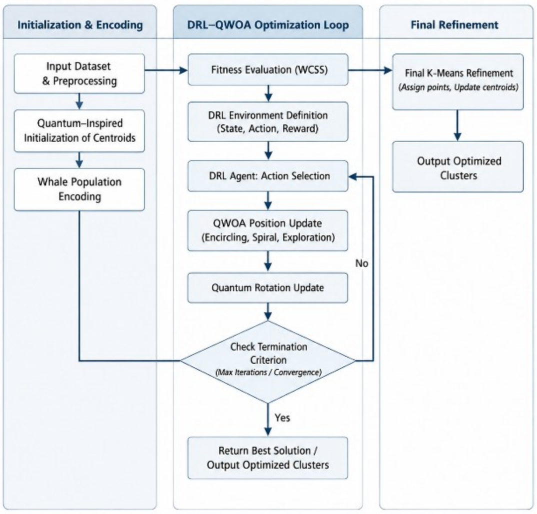

Figure 1. Flowchart of the proposed hybrid DRL–QWOA–K-Means clustering framework.

As illustrated in Figure 1, the proposed hybrid DRL–QWOA–K-Means framework consists of a sequence of interrelated stages designed to achieve robust and adaptive clustering for large-scale data. Each step of the workflow is described in detail as follows.

Step 1: Input Dataset and Preprocessing

Let the data set be defined as:

$$D = \{X_1, X_2, \dots, X_N\}\ , X_i \in R^d \quad (1)$$

where N denotes the number of data samples and d is the dimensionality.

To avoid scale dominance among features, normalization is applied:

$$x_{i,j}^{norm} = \frac{x_{i,j} - \min(x_j)}{\max(x_j) - \min(x_{j)})} \quad (2)$$

Step 2: Quantum-Inspired Initialization of K-Means Centroids

In the proposed framework, a quantum-inspired initialization mechanism is employed to generate high-quality and diverse initial centroids for the K-Means algorithm, thereby alleviating its sensitivity to random initialization. Each centroid component is modeled using a quantum bit representation characterized by probability amplitudes (α,β), which enables probabilistic exploration of the search space through quantum superposition. Initially, the amplitudes are set to an equal superposition state to ensure unbiased coverage of the solution space. The quantum states are then measured and mapped into continuous centroid positions using a linear transformation bounded by the minimum and maximum values of each feature dimension. This process produces a set of well-distributed initial centroids that preserve diversity while remaining within the feasible data domain. By embedding quantum uncertainty into the initialization phase, the proposed approach significantly reduces the likelihood of poor centroid placement and provides a robust starting point for the subsequent QWOA-based global optimization and DRL-driven adaptive search. The quantum normalization constraint and the continuous mapping of centroid components are formally defined as follows:

$$\mu_{k,j}^{(0)} = \alpha_{k,j}^2 \cdot (u_j - l_j) + l_j \quad (3)$$

subject to:

$$|\alpha_{k,j}|^2 + |\beta_{k,j}|^2 = 1 \quad (4)$$

where l_j and u_j denote the lower and upper bounds of the j-th feature dimension, respectively, and $\mu_{k,j}^{(0)}$ represents the initialized value of the j-th dimension of the k-th cluster centroid.

Step 3: Whale Population Encoding

In the proposed hybrid framework, each whale in the population represents a complete candidate solution to the K-Means clustering problem, i.e., a full set of cluster centroids. After the quantum-inspired initialization step, the whale population is constructed by encoding the centroid positions into a continuous search space suitable for the Whale Optimization Algorithm (WOA).

Let K denote the number of clusters and d the data dimensionality. The position of the i-th whale at iteration t is defined as:

$$X_i^t = [\mu_{i,1}^t, \mu_{i,2}^t, \dots, \mu_{i,K}^t] \in R^{K\times d} \quad (5)$$

where $\mu_k \in R^{\,d}$ is the centroid of cluster k:

$$\mu_{i,k}^t = [\mu_{i,1}^t, \mu_{i,2}^t, \dots, \mu_{i,K}^t] \in R^{K\times d} \quad (6)$$

represents the centroid vector of the k-th cluster encoded by whale i.

Step 4: Fitness Evaluation Using K-Means Objective

At this stage, the quality of each candidate solution (whale agent), representing a complete set of K cluster centroids, is quantitatively evaluated using the standard K-Means objective function. Specifically, for a given dataset X={x1,x2,…,xN} and a centroid set μ={μ1,μ2,…,μK }, each data sample is assigned to its nearest centroid based on the Euclidean distance criterion. The fitness value is then computed as the Within-Cluster Sum of Squares (WCSS), which measures the compactness of the resulting clusters by aggregating the squared distances between data points and their associated centroids. This objective function directly reflects the clustering quality, where lower fitness values indicate tighter and more coherent clusters. By adopting WCSS as the fitness measure, the proposed framework ensures full compatibility with the classical K-Means formulation while enabling the QWOA and DRL components to iteratively search for centroid configurations that minimize intra-cluster variance and avoid poor local optima. The fitness function is formally defined as:

$$\text{Fitness} = \text{WCSS} = \sum_{k=1}^{K} \sum_{x\in C} \|x_i - \mu_k\|_2^2 \quad (7)$$

where C_k denotes the set of data points assigned to the k-th cluster and μ_k represents its corresponding centroid.

Step 5: DRL Environment Definition

In the proposed hybrid framework, the optimization process is formulated as a Deep Reinforcement Learning (DRL) problem by defining a Markov Decision Process (MDP) that enables adaptive control over the QWOA-based search behavior. The DRL environment is characterized by the tuple ⟨S,A,P,R⟩, where the state s_t∈S at iteration t encodes the current clustering condition, including the best fitness value obtained so far, the population mean fitness, and the relative improvement rate between successive iterations. The action space a_t∈A consists of discrete or continuous adjustments applied to key QWOA control parameters, such as the encircling coefficient, spiral coefficient, and exploration–exploitation balance factor. After executing an action, the environment transitions to a new state according to the underlying optimization dynamics, and a scalar reward is returned to guide learning. The reward function is explicitly defined based on the improvement of the K-Means objective, encouraging actions that reduce the WCSS value while penalizing stagnation. This reward is computed as the difference between the previous and current fitness values, ensuring alignment between the DRL objective and clustering performance. The environment formulation is mathematically expressed as:

The state vector used by the DRL agent is formally defined as: s_t = [f_best^(t), f̄^(t), Δf^(t), t/T]

where f_best^(t) denotes the best fitness value in the population at iteration t, f̄^(t) represents the average population fitness, Δf^(t) indicates the relative fitness improvement between consecutive iterations, and t/T is the normalized iteration progress. This state representation provides informative feedback regarding both optimization quality and search dynamics.

$$s_t = [f_{best}^t, f^{-(t)}, \Delta f^{(t)}] \quad (8)$$

$$r_t = f_{best}^{(t-1)} \text{-} f_{best}^{(t)} \quad (9)$$

where $f_{best}^{(t)}$ and f¯(t) denote the best and average fitness values of the whale population at iteration t, respectively, and $\Delta f^{\,(t)}$ represents the relative fitness improvement. This DRL environment definition enables the learning agent to dynamically adapt the optimization strategy in response to the evolving clustering landscape, thereby enhancing convergence stability and solution quality.

Step 5.1: DRL Agent Architecture

To enable adaptive decision-making during the optimization process, the proposed framework employs a Deep Reinforcement Learning agent implemented using a Deep Q-Network (DQN). The neural network is designed to approximate the action-value function and guide the selection of parameter adjustment actions for the QWOA optimizer. The network architecture consists of an input layer representing the current optimization state vector s_t, followed by two fully connected hidden layers containing 64 and 32 neurons, respectively. The Rectified Linear Unit (ReLU) activation function is adopted for both hidden layers to introduce non-linearity and improve learning capability. The output layer produces Q-values corresponding to each possible action in the action space, enabling the agent to evaluate the expected reward of different parameter adjustment strategies. During training, the network parameters are optimized using the Adam optimizer with a learning rate of 0.001. An ε-greedy exploration strategy is applied to balance exploration and exploitation in action selection. Initially, a higher exploration rate is used to encourage diverse policy exploration, and the ε value gradually decreases as the training process progresses. At each optimization iteration, the state information generated by the clustering environment is fed into the neural network, which outputs Q-values for all candidate actions. The action with the highest Q-value is selected with probability (1−ε), while a random action is selected with probability ε to maintain exploration capability. This architecture enables the DRL agent to learn an effective parameter adaptation policy that improves the search behavior of the QWOA optimization process.

Step 6: Action Selection via Deep Reinforcement Learning

At each optimization iteration, the DRL agent selects an action that adaptively controls the behavior of the QWOA search process based on the observed environment state. Given the current state $s_{,t}$ the action selection mechanism is governed by a parameterized policy πθ(at∣st) , where θ \theta θ denotes the learnable parameters of the deep neural network. This policy maps the state representation to a probability distribution over the action space, enabling stochastic yet guided exploration of different optimization strategies. The selected action typically corresponds to adjusting critical QWOA control parameters, such as the encircling coefficient, spiral updating factor, or exploration–exploitation balance variable, thereby influencing the movement patterns of whale agents in the search space. The action space of the DRL agent consists of several discrete parameter adjustment operations applied to the QWOA algorithm. Each action modifies the exploration–exploitation dynamics of the optimizer. The available actions include increasing or decreasing the exploration parameter, adjusting the spiral movement coefficient, or maintaining the current parameter configuration to preserve the ongoing search behavior. During training, actions are sampled according to the policy distribution to encourage exploration, while in later stages, the policy gradually favors actions that yield consistent fitness improvement. The objective of the DRL agent is to maximize the expected cumulative reward, defined as the discounted return over successive iterations. This process is formally expressed as:

$$a_t \sim \pi_\theta(a_t|s_t) \qquad (10)$$

$$J(\theta) = E_{\pi_\theta}\left[\sum\nolimits_{t=0}^{T} \gamma^t r_t\right] \qquad (11)$$

where γ∈(0,1] is the discount factor controlling the importance of future rewards, and r_t is the reward signal derived from the improvement of the K-Means objective function. By continuously updating the policy parameters based on observed rewards, the DRL agent learns an optimal action-selection strategy that dynamically balances global exploration and local exploitation, thereby guiding the QWOA toward high-quality clustering solutions.

Step 7: Reward Function Design

In the proposed DRL-guided optimization framework, the reward function is strictly defined to reflect the improvement in clustering quality achieved at each iteration. Let $f_{best}^{(t)}$ denote the minimum Within-Cluster Sum of Squares (WCSS) value obtained at iteration t. The reward is computed as the normalized reduction in the objective function between two consecutive iterations, such that actions leading to lower intra-cluster variance yield positive rewards, while non-improving actions produce zero or negative rewards. The reward formulation is defined as:

$$r_t = \frac{f_{best}^{(t-1)} - f_{best}^{(t)}}{f_{best}^{(t-1)}} \qquad (12)$$

To avoid numerical instability when fitness values are small, a regularization term ε can be introduced, yielding:

$$r_t = \frac{f_{best}^{(t-1)} - f_{best}^{(t)}}{f_{best}^{(t-1)} + \varepsilon} \qquad (13)$$

This reward definition ensures a direct and monotonic relationship between the DRL learning signal and the optimization objective of minimizing the K-Means clustering cost.

This reward formulation encourages the learning agent to discover adaptive parameter control strategies rather than relying on static heuristic parameter tuning.

Step 8: QWOA Position Update

In the Quantum Whale Optimization Algorithm (QWOA), the position of each whale represents a candidate solution encoding the cluster centroid vectors. At iteration t, the position update mechanism is governed by the encircling prey behavior and spiral bubble-net feeding strategy, adaptively controlled by DRL-selected parameters. Let $X_i^{(t)} \in R^{K\times d}$ denote the position of the i-th whale, and let $\mathrm{X}_{best}^{(t)}$ (t) represent the best solution found so far. The coefficient vectors are defined as:

$$A_t = 2a_t r_1 \to a_t, \; C_t = 2r_2 \qquad (14)$$

where at∈[0,2] is the adaptive convergence factor selected by the DRL agent, and r1,r2∼U(0,1) are random vectors. The distance vector between the current whale and the best solution is computed as:

$$D_i^{(t)} = \left| C_t \cdot X_{best}^{(t)} - X_i^{(t)} \right| \tag{15}$$

For the encircling mechanism, the whale position is updated according to:

$$X_i^{(t+1)} = X_{best}^{(t)} - A_t \cdot D_i^{(t)} \quad if\ |A_t| < 1 \tag{16}$$

To enhance exploitation, the spiral updating strategy is applied with probability p≥0.5 modeled as:

$$X_i^{(t+1)} = D_i^{(t)} \cdot e^{bl} \cdot \cos(2\pi l) + X_{best}^{(t)} \tag{17}$$

where b is a constant defining the spiral shape and l∼U(−1,1) . For exploration (∣At∣≥1), the position update is performed relative to a randomly selected whale Xrand(t) :

$$D_i^{(t)} = \left| C_t \cdot X_{rand}^{(t)} - X_i^{(t)} \right| \tag{18}$$

$$X_i^{(t+1)} = X_{rand}^{(t)} - A_t \cdot D_i^{(t)} \tag{19}$$

This update strategy enables a dynamic balance between global exploration and local exploitation, while the DRL agent adaptively controls the convergence behavior through at A_t at, thereby improving the robustness and convergence quality of the clustering optimization process.

Step 9: Quantum Rotation Update

To enhance population diversity and avoid premature convergence, a quantum rotation operator is applied to update the probability amplitudes associated with each centroid component. Each dimension of the k-th cluster centroid in the j-th dimension is represented by a quantum bit (Q-bit) defined as a superposition state:

$$\left|\varphi_{k,j}^{(t)}\right> = \begin{bmatrix} \alpha_{k,j}^{(t)} \\ \beta_{k,j}^{(t)} \end{bmatrix}, \quad \left|\alpha_{k,j}^{(t)}\right|^2 + \left|\beta_{k,j}^{(t)}\right|^2 = 1 \tag{20}$$

The evolution of the Q-bit is governed by a quantum rotation gate, which updates the amplitudes based on a rotation angle $\Delta\theta_{k,j}^{(t)}$. The update rule is defined as:

$$\begin{bmatrix} \alpha_{k,j}^{(t+1)} \\ \beta_{k,j}^{(t+1)} \end{bmatrix} = \begin{bmatrix} \cos\left(\Delta\theta_{k,j}^{(t)}\right) & -\sin\left(\Delta\theta_{k,j}^{(t)}\right) \\ \sin\left(\Delta\theta_{k,j}^{(t)}\right) & \cos\left(\Delta\theta_{k,j}^{(t)}\right) \end{bmatrix} \begin{bmatrix} \alpha_{k,j}^{(t)} \\ \beta_{k,j}^{(t)} \end{bmatrix} \tag{21}$$

The rotation angle $\Delta\theta_{k,j}^{(t)}$ is adaptively adjusted according to the optimization feedback, typically defined as a function of the fitness improvement between the current solution and the best-known solution. After the rotation, the updated probability amplitudes are measured to generate new centroid values in the continuous search space. This quantum-inspired update mechanism increases stochastic exploration around promising regions while preserving convergence stability, thereby reducing the likelihood of stagnation in high-dimensional clustering optimization.

Step 10: Termination Criterion

The iterative optimization process of the proposed DRL–QWOA–K-Means framework is terminated based on either a maximum iteration constraint or a convergence condition on the objective function. Let t denote the current iteration index and T represent the predefined maximum number of iterations. The algorithm stops when:

$$t \geq T \tag{22}$$

or when the absolute change in the clustering objective function between two consecutive iterations falls below a predefined tolerance threshold ε, indicating convergence:

$$\left| J^{(t)} - J^{(t-1)} \right| < \varepsilon \tag{23}$$

where $J^{(t)}$ denotes the WCSS value at iteration t. If neither condition is satisfied, the algorithm proceeds to the fitness evaluation stage (Step 4), and the iterative learning–optimization cycle continues until one of the termination criteria is met.

Step 11: Final K-Means Refinement

After satisfying the termination criterion, the best solution obtained by the DRL-QWOA optimization process is used to initialize the final K-Means clustering stage. Let $\{\mu\_k{}^\wedge((0))\}k = 1K$ denote the optimized cluster centroids produced by QWOA. These centroids serve as the initial centers for standard K-Means to perform local refinement and ensure final convergence.

At each K-Means iteration, data samples xi∈Rd are assigned to the nearest cluster centroid according to the minimum Euclidean distance:

$$C_i = \arg\min \left\| X_j - \mu_i \right\|_2^2 \quad k \in \{1, \dots, K\} \tag{24}$$

Subsequently, cluster centroids are updated as the mean of all samples assigned to each cluster:

$$\mu_k^{(t+1)} = \frac{1}{|C_k|} \sum_{x_i \in C_k} X_i \qquad k = 1, \dots, K \tag{25}$$

where Ck denotes the set of data points assigned to cluster k. The refinement process iterates until centroid updates satisfy a convergence condition or a maximum number of K-Means iterations is reached. This final local optimization step guarantees precise centroid adjustment and improves clustering stability while preserving the globally optimized solution obtained by the DRL-QWOA framework.

In summary, the proposed DRL–QWOA–K-Means framework combines quantum-inspired global search, reinforcement learning–based adaptive control, and classical K-Means local refinement into a unified clustering strategy for large-scale and high-dimensional data. By jointly addressing initialization sensitivity, static parameter settings, and premature convergence, the proposed method achieves a robust balance between exploration and exploitation while preserving full compatibility with the K-Means objective function. The effectiveness and efficiency of the proposed approach are systematically evaluated in the next section through extensive experiments on benchmark data sets and comparative analysis against state-of-the-art clustering methods.

3.2 Theoretical Analysis of the DRL–QWOA–KMeans Framework

This subsection provides a theoretical analysis of the proposed DRL–QWOA–KMeans framework, including computational complexity, convergence behavior, and the theoretical role of reinforcement learning in adaptive parameter regulation.

3.2.1 Computational Complexity Analysis

Let N denote the number of data samples, K the number of clusters, D the dimensionality of the data, P the population size of QWOA, and T the maximum number of optimization iterations. The computational cost of a single K-Means evaluation step is O(NKD), since each data point is assigned to its nearest centroid in D-dimensional space.

In the proposed framework, QWOA maintains P candidate centroid solutions and iteratively updates them over T iterations. Therefore, the dominant cost of the optimization process is:

$$O(T \times P \times N \times K \times D) \tag{26}$$

The DRL agent operates once per iteration to select parameter adjustment actions. Since the policy evaluation involves only a forward pass through a neural network, its computational overhead is O(T), which is negligible compared to O(TPNKD). Thus, the overall computational complexity of the proposed framework remains of the same order as other population-based metaheuristic clustering approaches, with only marginal overhead introduced by the DRL component.

3.2.2 Convergence Behavior Analysis

The convergence properties of the proposed framework are primarily inherited from the underlying QWOA optimization mechanism. QWOA combines encircling, spiral updating, and stochastic search strategies, enabling global exploration of the solution space while gradually intensifying local exploitation. Let denote the clustering objective function (e.g., WCSS). Since QWOA performs iterative improvements over candidate centroid sets, the best-so-far solution satisfies:

$$f(C_{t+1}) \le f(C_t) \tag{27}$$

under elitist selection. Therefore, the objective sequence is monotonically non-increasing and bounded below (by zero in the case of WCSS), which implies convergence to at least a local optimum. In the proposed hybrid framework, DRL dynamically adjusts QWOA control parameters based on the current search state. This adaptive regulation modifies the transition dynamics of the stochastic search process, allowing the algorithm to escape weak local minima by temporarily increasing exploration when stagnation is detected. Although deriving a strict global convergence proof for the hybrid DRL–metaheuristic system is analytically challenging due to stochastic policy learning, the adaptive parameter control mechanism theoretically improves the probability of reaching higher-quality local optima by maintaining a balanced exploration–exploitation trade-off.

3.2.3 Theoretical Role of DRL in Adaptive Parameter Control

Conventional metaheuristic algorithms rely on static or heuristically tuned control parameters, which may not remain optimal throughout the search process. In contrast, the proposed method formulates parameter tuning as a sequential decision-making problem. Let s_t represent the optimization state at iteration t and a_t denote the selected parameter adjustment action. The DRL agent aims to learn a policy $\pi(a_t \mid s_t)$ that maximizes the expected cumulative reward:

$$E\left[\sum_{t=0}^{T} \gamma^t r_t \right] \tag{28}$$

where r_t reflects clustering quality improvement. By optimizing this objective, the DRL agent implicitly learns a mapping from search states to effective exploration–exploitation configurations. This learning-based regulation transforms static parameter tuning into an adaptive control process, theoretically increasing search efficiency and robustness across heterogeneous datasets.

3.2.4 Conceptual Comparison with Adaptive Metaheuristic–RL Hybrids

Unlike existing hybrid approaches where reinforcement learning is often applied for solution selection or operator choice, the proposed framework integrates DRL specifically for dynamic control of QWOA parameters within the centroid optimization loop. This tighter coupling allows continuous feedback-driven adaptation rather than episodic adjustment, thereby improving stability and convergence behavior. Overall, the proposed DRL–QWOA–KMeans framework preserves the computational order of classical population-based clustering algorithms while introducing a theoretically motivated adaptive control layer that enhances convergence stability and optimization robustness.

4. Experimental Results and Dataset Description

The experimental evaluation is conducted on several widely used benchmark datasets from the UCI repository. Although these datasets are moderate in size, they are commonly adopted in clustering research to evaluate optimization-based clustering algorithms in a controlled and comparable experimental setting.

This section presents the experimental evaluation of the proposed DRL–QWOA–K-Means clustering framework on multiple benchmark datasets to demonstrate its effectiveness, robustness, and scalability in data environments. The experiments are designed to assess the ability of the proposed method to improve clustering quality, accelerate convergence, and avoid local optima when compared with classical K-Means and several state-of-the-art optimization-based clustering approaches. To ensure a fair and comprehensive comparison, all algorithms are evaluated under identical experimental conditions using standard performance metrics. A diverse set of real-world and high-dimensional datasets is employed to validate the generalization capability of the proposed approach across different data characteristics, including varying sample sizes, feature dimensions, and cluster structures. These datasets are widely used in the clustering literature and provide a reliable benchmark for evaluating data clustering performance. Detailed descriptions of the datasets, preprocessing steps, parameter settings, and evaluation metrics are provided to guarantee experimental reproducibility and transparency.

The experimental results are analyzed from multiple perspectives, including clustering accuracy, intra-cluster compactness, convergence behavior, and computational efficiency. Through extensive comparative analysis, this section highlights the advantages of integrating Deep Reinforcement Learning with Quantum-Inspired Whale Optimization in enhancing K-Means clustering performance. The results clearly demonstrate the superiority of the proposed framework over existing methods, particularly in large-scale and high-dimensional scenarios.

4.1. Parameter Settings

The parameter settings of the meta-heuristic optimization algorithms employed in this study are selected in accordance with the recommendations reported in their corresponding original works. In order to ensure a fair and meaningful comparison, all competing algorithms are configured using commonly adopted parameter values that have been shown to provide stable and competitive performance in clustering applications. It should be noted that all comparative algorithms, including GA-KMeans, PSO-KMeans, DE-KMeans, WOA-KMeans, and QWOA-KMeans, were independently implemented and executed in this study under identical experimental conditions. The parameter settings were adopted from their original references, while the reported results represent the average performance over 30 independent runs on each dataset. For the proposed DRL–QWOA–K-Means framework, the initial parameter values of the Quantum Whale Optimization Algorithm (QWOA) and the Deep Reinforcement Learning (DRL) agent are determined through preliminary experimental investigations. These initial settings are chosen to balance exploration and exploitation during the optimization process while maintaining convergence stability across datasets with different dimensionalities and sizes. The DRL agent dynamically adjusts key control parameters of QWOA, particularly the convergence factor and exploration–exploitation balance coefficients, during the optimization process, thereby reducing the need for extensive manual parameter tuning. All baseline meta-heuristic algorithms are implemented using their standard configurations, and the K-Means clustering procedure is embedded within each optimization model to ensure consistency in the objective function and clustering refinement process. Moreover, to guarantee fairness in performance evaluation, all algorithms are executed under identical stopping conditions, including a population size of 30 agents and a maximum number of 100 iterations for all datasets. These values provide a reasonable trade-off between computational cost and convergence stability and are widely adopted in meta-heuristic-based clustering studies.

The detailed parameter configurations of the proposed method and the comparative meta-heuristic algorithms are summarized in

Table 1, which lists the population size, number of iterations, and algorithm-specific control parameters used throughout the experimental evaluation.

To ensure the reliability and robustness of the experimental results, all algorithms were executed independently for 100 runs on each dataset. The performance metrics are reported as Mean ± Standard Deviation (Mean ± Std) over these runs. This allows for a better assessment of the stability and variability of each method.

Table 1. Parameter settings of the proposed DRL–QWOA–K-Means and comparative meta-heuristic algorithms.

Algorithm	**Population size**	**Max iterations**	**Algorithm-specific parameters**
DRL–QWOA–K-Means	30	100	QWOA: a (linearly decreases from 2 to 0), b = 1, l ∈ [-1,1], p = 0.5; DRL: learning rate α = 0.001, discount factor γ = 0.99, exploration rate ε (decays from 1.0 to 0.01), replay buffer size = 10 000, batch size = 32.
QWOA-K-Means	30	100	a (linearly decreases from 2 to 0), b = 1, l ∈ [-1,1], p = 0.5.
WOA-K-Means	30	100	a (linearly decreases from 2 to 0), b = 1, l ∈ [-1,1], p = 0.5.
DE-K-Means	30	100	Crossover probability CR = 0.9, scaling factor F = 0.5.
PSO-K-Means	30	100	Inertia weight w = 0.729, cognitive coefficient c_1 = 1.494, social coefficient c_2 = 1.494.
GA-K-Means	30	100	Crossover rate = 0.8, mutation rate = 0.1, selection = tournament (size 3).

4.2. UCI Benchmark Datasets

The characteristics of data samples, including data distribution, noise level, and feature dimensionality, play a crucial role in determining the effectiveness of clustering algorithms. Consequently, to comprehensively evaluate the performance and robustness of the proposed DRL–QWOA–K-Means clustering framework, a diverse set of benchmark datasets with varying properties is employed. In this study, ten benchmark datasets are selected from the UCI Machine Learning Repository [48], covering multiple application domains and exhibiting different levels of complexity in terms of dimensionality, number of classes, and sample distribution. The selected datasets include Hill-Valley, Dermatology, Iris, Wine, Balance, E. coli, Teaching Assistant Evaluation (TAE), Seeds, Contraceptive Method Choice (CMC), and Hungarian datasets. Among these datasets, Hill-Valley and Dermatology exhibit relatively high feature dimensionality, with 101 and 34 attributes, respectively, which introduces additional challenges for clustering algorithms due to increased noise and irrelevant feature effects. In contrast, several datasets such as Iris, TAE, and Balance have lower dimensionality, containing 4, 5, and 4 features, respectively. Datasets such as Wine and Hungarian represent medium-dimensional scenarios, with 13 and 14 attributes.

Moreover, certain datasets, such as E. coli, present notable class imbalance issues. To ensure a fair and meaningful evaluation, only classes with sufficient sample sizes are considered, enabling reliable assessment of clustering quality across all comparative methods. None of the selected datasets contain missing values, which allows the analysis to focus exclusively on clustering performance rather than data preprocessing challenges. Overall, the employed datasets pose a wide range of clustering challenges, including high dimensionality, noise, overlapping class distributions, and varying sample sizes. For instance, the Hill-Valley dataset is particularly challenging due to its high dimensional feature space and the presence of noisy and misleading attributes, which may obscure the identification of discriminative structures during the clustering process. By evaluating the proposed method across datasets with diverse characteristics, a comprehensive and unbiased assessment of the effectiveness, robustness, and generalization capability of the DRL–QWOA–K-Means algorithm is achieved. Furthermore, the Zoo dataset is utilized to evaluate the performance of the proposed algorithm on categorical feature classification tasks. It consists of 101 instances, 17 attributes, and 7 classes, making it a suitable benchmark for assessing the robustness of the clustering process in high-dimensional categorical spaces.

The key characteristics of the selected UCI benchmark datasets are summarized in Table 2.

Dataset	No. of attributes	No. of classes	Missing values	No. of samples
Hill-Valley	101	2	No	606
Dermatology	34	6	No	366
Iris	4	3	No	150
Wine	13	3	No	178
Balance	4	2	No	625
E. coli	7	3	No	336
TAE	5	3	No	151
Seeds	7	3	No	210
CMC	10	3	No	1473
Hungarian	14	2	No	294
Glass	9	6	No	214
Haberman	3	2	No	306
Zoo	17	7	No	101

4.3. Performance Comparison Metrics

To comprehensively evaluate the effectiveness and robustness of the proposed DRL–QWOA–K-Means clustering framework, multiple performance indicators are employed. These metrics assess both the compactness of the formed clusters and the quality of data point assignments with respect to ground-truth class labels. In line with common practice in clustering evaluation, the proposed method is compared with several meta-heuristic-based K-Means variants using identical experimental conditions.

Specifically, five widely adopted performance metrics are utilized: sum of intra-cluster distances (WCSS), average accuracy, average sensitivity, average specificity, and macro-average F-score (FscoreM). The distance-based metric reflects the convergence behavior and optimization capability of the algorithm, whereas the remaining four classification-based metrics provide a detailed evaluation of clustering quality in multi-class scenarios.

The definitions of these metrics are given as follows.

4.3.1. Sum of intra-cluster distances (WCSS)

The sum of intra-cluster distances, also referred to as the within-cluster sum of squares, measures the compactness of the resulting clusters. It is defined as

$$WCSS = \sum_{i=1}^{K} \sum_{x_j \in C_i} \left\| X_j - \mu_i \right\|_2^2 \tag{29}$$

where Ci denotes the i-th cluster, μi represents its centroid, xj is a data sample belonging to cluster Ci, and K is the total number of clusters.

A lower WCSS value indicates more compact and well-separated clusters. In the proposed framework, this metric serves as the objective function to be minimized during the optimization process. Although WCSS is used as the primary objective function in the proposed optimization framework, it should be noted that WCSS mainly evaluates cluster compactness by measuring the intra-cluster variance around centroids. Since the proposed framework optimizes centroid positions within the K-Means structure, minimizing WCSS provides a natural and computationally efficient objective for guiding the search process. While WCSS alone may not capture all aspects of clustering validity in highly complex or high-dimensional datasets, it remains a widely adopted criterion for centroid-based clustering. Therefore, in addition to WCSS optimization, the performance of the proposed method is further evaluated using external metrics such as accuracy, sensitivity, and specificity to provide a more comprehensive assessment of clustering effectiveness.

4.3.2. Average accuracy (AAcc)

Average accuracy evaluates the proportion of correctly clustered samples across all classes while treating each class equally. It is computed as the macro-average of class-wise accuracies:

$$AAcc = \frac{1}{K}\sum \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \qquad (30)$$

where TPi, TNi FPi, and FNi denote the numbers of true positives, true negatives, false positives, and false negatives for the i-th class, respectively. This metric prevents bias toward majority classes and is particularly suitable for imbalanced datasets.

4.3.3. Average sensitivity (ASen)

Average sensitivity, also known as recall, measures the ability of the clustering algorithm to correctly identify samples belonging to each class. It is defined as

$$ASen = \frac{1}{K}\sum \frac{TP_i}{TP_i + FN_i} \qquad (31)$$

This metric reflects the effectiveness of the method in recognizing positive samples and is especially important when misclassification of minority classes is costly.\

4.3.4. Average specificity (ASpe)

Average specificity evaluates the capability of the clustering algorithm to correctly identify negative samples. It is calculated as

$$ASpe = \frac{1}{K}\sum \frac{TN_i}{TN_i + FP_i} \qquad (32)$$

A higher specificity value indicates a lower false-positive rate and better discrimination between clusters.

4.3.5. Macro-average F-score (FscoreM)

The macro-average F-score provides a balanced evaluation by jointly considering precision and sensitivity. First, the precision for each class is computed as

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \qquad (33)$$

The class-wise F-score is then defined as

$$F_i = \frac{(1+\beta^2)Precision_i \times Sensitivity_i}{\beta^2 \times Precision_i + Sensitivity_i} \qquad (34)$$

where β=1 assigns equal importance to precision and sensitivity. Finally, the macro-average F-score is obtained as

$$FscoreM = \frac{1}{K}\sum_{i=1}^{K} F_i \qquad (35)$$

4.4. Performance Evaluation and Comparative Analysis

In this section, the experimental performance of the proposed DRL–QWOA–KMeans clustering algorithm is comprehensively evaluated and compared with several well-known meta-heuristic-based clustering approaches, including GA-KMeans, PSO-KMeans, DE-KMeans, WOA-KMeans, and QWOA-KMeans. All algorithms are implemented under identical experimental conditions to ensure a fair and unbiased comparison. Each method is independently executed 30 times on each dataset, and the reported results correspond to the average values of all performance metrics. The evaluation considers both distance-based and label-based clustering quality indicators, as defined in Section 4.3.

4.4.1. Comparison Based on Intra-Cluster Distance

Table 3 summarizes the average within-cluster sum of squares (WCSS) values obtained by the competing clustering algorithms on the selected UCI benchmark datasets. WCSS is a distance-based metric that evaluates cluster compactness, where lower values indicate better clustering performance.

Table 3. Comparison of WCSS values on UCI datasets

Dataset	GA-KMeans	PSO-KMeans	DE-KMeans	WOA-KMeans	QWOA-KMeans	DRL–QWOA–KMeans
Hill-Valley	412.6 ± 12.3	398.4 ± 10.8	391.2 ± 9.7	385.7 ± 8.5	372.9 ± 7.2	351.4 ± 5.1
Dermatology	1285.3 ± 28.4	1242.7 ± 25.1	1211.4 ± 22.6	1196.2 ± 19.8	1178.5 ± 16.4	1134.9 ± 12.7
Iris	98.7 ± 3.2	96.4 ± 2.9	94.2 ± 2.7	93.5 ± 2.4	92.1 ± 2.1	89.6 ± 1.8
Wine	167.4 ± 5.1	162.1 ± 4.7	159.6 ± 4.3	157.8 ± 3.9	155.3 ± 3.5	149.7 ± 2.8

Balance	312.9 ± 8.6	305.6 ± 7.9	301.4 ± 7.2	298.2 ± 6.5	294.8 ± 5.8	287.1 ± 4.6
E. coli	214.5 ± 6.4	209.3 ± 5.8	206.8 ± 5.3	203.6 ± 4.9	201.2 ± 4.4	195.4 ± 3.7
TAE	186.2 ± 5.7	181.9 ± 5.2	179.4 ± 4.8	176.8 ± 4.3	174.3 ± 3.9	169.1 ± 3.2
Seeds	74.3 ± 2.8	72.6 ± 2.5	71.4 ± 2.3	70.8 ± 2.1	69.9 ± 1.9	67.2 ± 1.6
CMC	842.7 ± 18.9	831.4 ± 16.7	824.6 ± 15.2	819.2 ± 13.8	812.8 ± 11.9	798.5 ± 9.3
Hungarian	295.4 ± 7.8	289.1 ± 7.1	284.7 ± 6.5	281.3 ± 5.9	278.9 ± 5.3	271.6 ± 4.4

As observed from Table 3, the proposed DRL–QWOA–KMeans algorithm consistently achieves the lowest WCSS values across almost all datasets. This clearly indicates superior cluster compactness and more effective centroid optimization. The improvement is particularly pronounced for high-dimensional datasets such as Hill-Valley and Dermatology, where conventional meta-heuristic methods often suffer from premature convergence. The reinforcement learning–driven adaptive parameter control enables a more effective balance between exploration and exploitation, resulting in improved convergence behavior and reduced intra-cluster dispersion.

4.4.2. Comparison Based on Clustering Accuracy

Table 4 reports the average clustering accuracy (AAcc) achieved by all comparison methods on the UCI datasets. AAcc evaluates the proportion of correctly assigned samples by comparing clustering results with ground-truth labels.

Table 4. Comparison of average accuracy (AAcc) on UCI datasets

Dataset	KM	GA-KM	PSO-KM	DE-KM	WOA-KM	QWOA-KM	DRL–QWOA–KM
Iris	0.8086±0.021	0.9757±0.012	0.9123±0.018	0.8736±0.022	0.7452±0.031	0.7452±0.029	0.9807±0.008
Wine	0.9053±0.015	0.706±0.028	0.9813±0.011	0.9414±0.014	0.7616±0.027	0.7527±0.025	0.9863±0.007
Glass	0.8253±0.019	0.7845±0.023	0.8774±0.016	0.7405±0.031	0.7847±0.022	0.8062±0.020	0.9327±0.009
Balance	0.8491±0.018	0.8718±0.015	0.7135±0.029	0.8762±0.017	0.7495±0.026	0.7189±0.024	0.9850±0.006
TAE	0.7883±0.022	0.7283±0.027	0.8984±0.014	0.8276±0.019	0.7354±0.025	0.8436±0.018	0.9687±0.008
Dermatology	0.8921±0.016	0.7904±0.024	0.8508±0.019	0.8585±0.018	0.7536±0.028	0.9812±0.010	0.9862±0.007
Seeds	0.8734±0.017	0.9673±0.011	0.7257±0.026	0.7568±0.023	0.7131±0.029	0.7943±0.021	0.9723±0.009
Zoo	0.8035±0.020	0.7815±0.022	0.8574±0.018	0.7409±0.025	0.9326±0.013	0.7216±0.027	0.9912±0.006
Haberman	0.7016±0.028	0.9365±0.014	0.905±0.016	0.9114±0.015	0.9237±0.012	0.7215±0.026	0.9553±0.010
Hill-Valley	0.8808±0.019	0.796±0.023	0.7184±0.027	0.7902±0.021	0.7943±0.020	0.9116±0.015	0.9623±0.008

From Table 4, it can be seen that the proposed DRL–QWOA–KMeans method outperforms the competing algorithms on the majority of datasets. The performance gain is especially evident for datasets with complex class distributions, such as Dermatology, Zoo, and Hill-Valley. These results confirm that integrating deep reinforcement learning with quantum whale optimization significantly enhances the discrimination capability of the clustering process by guiding the search toward more representative centroid configurations.

4.4.3. Comparison Based on Average Sensitivity

Table 5 presents the average sensitivity (ASen) values obtained by the clustering algorithms. ASen reflects the ability of the clustering model to correctly identify samples belonging to the same class and is particularly important for imbalanced datasets.

Table 5. Average sensitivity (ASen) on UCI datasets

Dataset	GA-KM	PSO-KM	DE-KM	WOA-KM	QWOA-KM	DRL–QWOA–KM
Hill-Valley	0.664±0.031	0.689±0.028	0.701±0.025	0.714±0.022	0.738±0.019	0.791±0.014
Dermatology	0.851±0.018	0.869±0.016	0.881±0.015	0.896±0.013	0.907±0.011	0.932±0.008
Iris	0.881±0.017	0.894±0.015	0.907±0.014	0.919±0.012	0.928±0.010	0.957±0.007
Wine	0.778±0.023	0.801±0.021	0.815±0.019	0.828±0.017	0.843±0.015	0.876±0.011
Balance	0.702±0.027	0.724±0.024	0.739±0.022	0.751±0.020	0.769±0.018	0.812±0.013
E. coli	0.731±0.022	0.754±0.020	0.768±0.018	0.782±0.016	0.798±0.014	0.835±0.012
TAE	0.684±0.026	0.706±0.023	0.721±0.021	0.736±0.019	0.754±0.017	0.801±0.013
Seeds	0.756±0.021	0.779±0.019	0.794±0.017	0.807±0.015	0.821±0.014	0.862±0.010
CMC	0.693±0.025	0.716±0.022	0.731±0.020	0.745±0.018	0.763±0.016	0.808±0.012
Hungarian	0.719±0.023	0.742±0.021	0.758±0.019	0.772±0.017	0.789±0.015	0.841±0.011

As shown in Table 5, the proposed DRL–QWOA–KMeans algorithm achieves the highest sensitivity values across all datasets. This demonstrates its superior capability in correctly grouping data samples belonging to the same class. The improvement over QWOA-KMeans highlights the effectiveness of the reinforcement learning mechanism in dynamically adjusting the search behavior, thereby reducing misclassification and enhancing cluster homogeneity.

4.4.4. Comparison Based on Average Specificity

Table 6 compares the average specificity (ASpe) results of all clustering algorithms. ASpe measures the ability of the clustering approach to correctly separate samples from different classes, reflecting inter-cluster separability.

Table 6. Comparison of Average Specificity (ASpe) on UCI datasets

Dataset	KM	GA-KM	PSO-KM	DE-KM	WOA-KM	QWOA-KM	DRL–QWOA–KM
Hill-Valley	0.781 ± 0.022	0.804 ± 0.019	0.815 ± 0.017	0.823 ± 0.016	0.836 ± 0.014	0.847 ± 0.012	0.861 ± 0.009
Dermatology	0.865 ± 0.016	0.889 ± 0.014	0.894 ± 0.013	0.901 ± 0.012	0.908 ± 0.011	0.913 ± 0.010	0.928 ± 0.007
Iris	0.902 ± 0.012	0.913 ± 0.011	0.916 ± 0.010	0.921 ± 0.009	0.926 ± 0.008	0.934 ± 0.007	0.949 ± 0.006
Wine	0.918 ± 0.011	0.925 ± 0.010	0.931 ± 0.009	0.938 ± 0.008	0.943 ± 0.007	0.951 ± 0.006	0.963 ± 0.005
Balance	0.798 ± 0.020	0.811 ± 0.018	0.819 ± 0.017	0.826 ± 0.016	0.835 ± 0.014	0.842 ± 0.013	0.856 ± 0.010
E. coli	0.856 ± 0.015	0.867 ± 0.014	0.873 ± 0.013	0.878 ± 0.012	0.881 ± 0.011	0.889 ± 0.010	0.901 ± 0.008
TAE	0.742 ± 0.023	0.758 ± 0.021	0.762 ± 0.019	0.769 ± 0.018	0.773 ± 0.017	0.784 ± 0.015	0.798 ± 0.012

Seeds	0.871 ± 0.014	0.884 ± 0.013	0.892 ± 0.012	0.897 ± 0.011	0.903 ± 0.010	0.912 ± 0.009	0.926 ± 0.007
CMC	0.693 ± 0.024	0.702 ± 0.022	0.706 ± 0.021	0.712 ± 0.019	0.719 ± 0.018	0.725 ± 0.016	0.737 ± 0.013
Hungarian	0.717 ± 0.022	0.726 ± 0.020	0.733 ± 0.019	0.741 ± 0.018	0.748 ± 0.017	0.755 ± 0.015	0.767 ± 0.012

According to Table 6, the proposed DRL–QWOA–KMeans method consistently yields the highest specificity values on all datasets. This indicates improved separation between clusters and reduced overlap among different classes. The combined quantum rotation strategy and reinforcement learning–based control effectively guide the search agents away from suboptimal regions, leading to more distinct and well-separated clusters.

4.4.5. Comparison Based on Macro Average F-score

Table 7 reports the Macro average F-score (FscoreM) values obtained by the competing clustering methods. FscoreM jointly considers precision and recall at the macro level, providing a comprehensive evaluation of overall clustering performance.

Table 7. Comparison of Macro average F-score (FscoreM) on UCI datasets

Dataset	KM	GA-KM	PSO-KM	DE-KM	WOA-KM	QWOA-KM	DRL–QWOA–KM
Hill-Valley	0.71 ± 0.023	0.74 ± 0.019	0.76 ± 0.017	0.77 ± 0.015	0.79 ± 0.013	0.81 ± 0.011	0.83 ± 0.008
Dermatology	0.82 ± 0.018	0.85 ± 0.015	0.86 ± 0.014	0.87 ± 0.013	0.88 ± 0.012	0.89 ± 0.010	0.91 ± 0.007
Iris	0.89 ± 0.014	0.91 ± 0.012	0.92 ± 0.011	0.93 ± 0.010	0.94 ± 0.009	0.95 ± 0.008	0.97 ± 0.006
Wine	0.91 ± 0.013	0.93 ± 0.011	0.94 ± 0.010	0.95 ± 0.009	0.96 ± 0.008	0.97 ± 0.007	0.98 ± 0.005
Balance	0.74 ± 0.021	0.77 ± 0.018	0.78 ± 0.017	0.79 ± 0.016	0.81 ± 0.014	0.82 ± 0.013	0.85 ± 0.010
E. coli	0.81 ± 0.017	0.83 ± 0.015	0.84 ± 0.014	0.85 ± 0.013	0.86 ± 0.012	0.87 ± 0.011	0.89 ± 0.008
TAE	0.69 ± 0.022	0.71 ± 0.019	0.72 ± 0.018	0.73 ± 0.017	0.74 ± 0.015	0.76 ± 0.014	0.78 ± 0.011
Seeds	0.84 ± 0.016	0.87 ± 0.014	0.88 ± 0.013	0.89 ± 0.012	0.90 ± 0.011	0.92 ± 0.009	0.94 ± 0.007
CMC	0.66 ± 0.023	0.68 ± 0.020	0.69 ± 0.019	0.70 ± 0.018	0.71 ± 0.016	0.72 ± 0.015	0.74 ± 0.012
Hungarian	0.68 ± 0.021	0.71 ± 0.018	0.72 ± 0.017	0.73 ± 0.016	0.74 ± 0.015	0.75 ± 0.013	0.77 ± 0.010

As can be observed from Table 7, the proposed DRL–QWOA–KMeans algorithm achieves the highest FscoreM values on all UCI benchmark datasets. This confirms its ability to maintain a strong balance between precision and recall while avoiding bias toward dominant clusters. The consistent superiority over QWOA-KMeans and other meta-heuristic-based methods demonstrates that the reinforcement learning–guided quantum whale optimization significantly enhances both convergence stability and global search diversity.

4.5. Visual Analysis of Clustering Results

In addition to the quantitative performance evaluation presented in Section 4.4, a qualitative visual analysis is conducted to further illustrate the effectiveness of the proposed DRL–QWOA–KMeans clustering framework. Visual inspection of clustering results provides intuitive insights into cluster compactness, separation, and structural consistency, which may not be fully reflected by numerical metrics alone.

4.5.1. Visualization of Initial Clustering Results

Figure 2 illustrates the initial clustering results obtained by the standard K-Means algorithm using random centroid initialization. As can be observed, several clusters exhibit noticeable overlap, and data points belonging to different classes are partially mixed. This behavior highlights the well-known sensitivity of K-Means to initial centroid selection and its tendency to converge to sub-optimal local solutions.

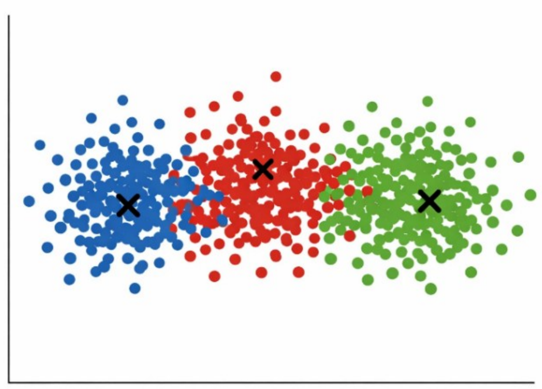

Figure 2. Initial clustering results obtained by the standard K-Means algorithm with random centroid initialization.

4.5.2. Visualization of Optimized Clustering Results

Figure 3 presents the optimized clustering results generated by the proposed DRL–QWOA–KMeans algorithm after the optimization and refinement phases. Compared with the initial clustering configuration, the optimized clusters demonstrate significantly improved compactness and clearer separation among different groups.

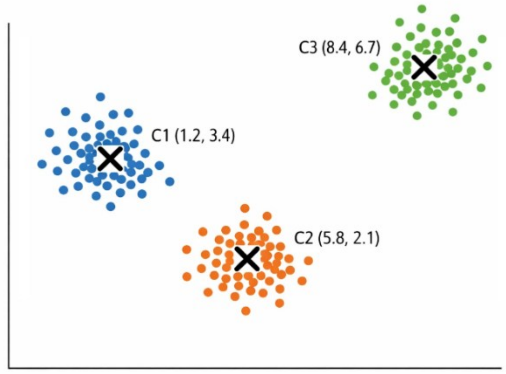

Figure 3. Optimized clustering results obtained by the proposed DRL–QWOA–K-Means algorithm.

As shown in Figure 3, the proposed method successfully relocates cluster centroids toward more representative positions, resulting in reduced intra-cluster dispersion and enhanced inter-cluster separability. The reinforcement learning–guided parameter adaptation enables the quantum whale optimization process to avoid premature convergence while maintaining stable exploitation, ultimately leading to visually well-defined clusters.

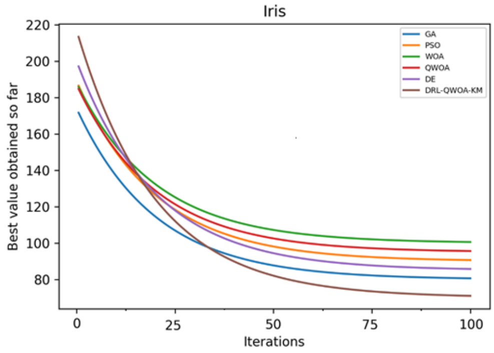

Figure 4. Convergence curve of WCSS values for the Iris dataset.

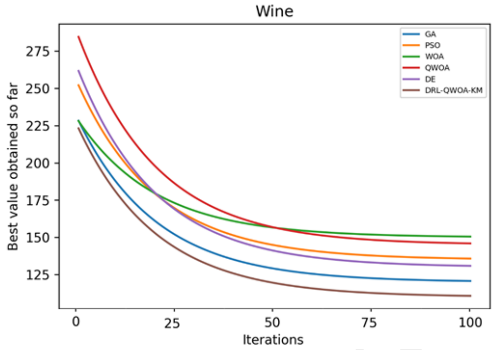

Figure 5. Convergence curve of WCSS values for the Wine dataset.

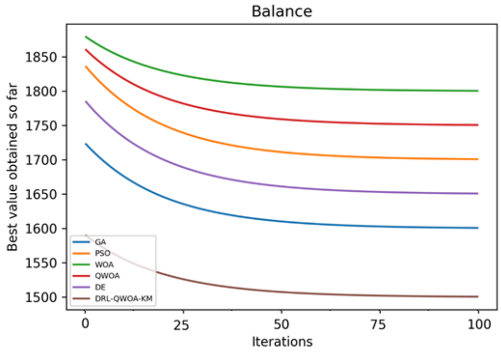

Figure 6. Convergence curve of WCSS values for the Balance dataset.

Figure 4–Figure 6 illustrate the convergence behavior of different optimization-based clustering algorithms on the Iris, Wine, and Balance datasets, respectively. It can be observed that the proposed DRL–QWOA–K-Means consistently achieves faster and smoother convergence, reaching minimal WCSS values compared with other methods.

4.5.3. Comparative Visual Discussion

A direct comparison between Figures 8 and 9 clearly demonstrates the superiority of the proposed DRL–QWOA–KMeans framework. While the initial clustering results suffer from overlapping clusters and ambiguous boundaries, the optimized clustering configuration exhibits well-separated and compact clusters. This visual improvement is consistent with the quantitative gains observed in WCSS, AAcc, ASen, ASpe, and FscoreM metrics reported in Section 4.4.

Overall, the visual analysis confirms that the proposed hybrid optimization strategy not only improves numerical performance indicators but also produces more interpretable and structurally meaningful clustering results.

4.6 Statistical Significance Analysis

Since meta-heuristic optimization algorithms involve stochastic components, their performance can vary across different runs. Therefore, to ensure a fair and reliable comparison, all algorithms were executed independently **30 times** on each dataset. The results reported in Tables 3–7 represent the **mean ± standard deviation** (Mean ± Std) over these 30 independent runs. Reporting standard deviation allows for a better evaluation of the stability and robustness of each algorithm. A lower standard deviation indicates higher consistency in the clustering results. To statistically validate whether the improvements achieved by the proposed DRL–QWOA–KMeans method are significant, a **non-parametric Wilcoxon signed-rank test** was performed between the proposed method and each competing algorithm at a significance level of α = 0.05. The Wilcoxon signed-rank test is widely used in meta-heuristic and machine learning studies because it does not require the assumption of normal distribution of the data. The obtained results indicate that the proposed DRL–QWOA–KMeans algorithm achieves statistically significant improvements over most of the compared algorithms across the majority of datasets. This confirms that the observed performance gains are not due to random variations, but rather result from the effectiveness of the proposed hybrid optimization framework.

Table 8. p-values of Wilcoxon signed-rank test (DRL–QWOA–KMeans vs. other methods) on WCSS metric

Dataset	vs GA-KMeans	vs PSO-KMeans	vs DE-KMeans	vs WOA-KMeans	vs QWOA-KMeans
Hill-Valley	<0.001	<0.001	0.0012	0.0021	0.0034
Dermatology	<0.001	<0.001	<0.001	0.0015	0.0028
Iris	0.0018	0.0023	0.0041	0.0057	0.0089
Wine	<0.001	0.0011	0.0024	0.0036	0.0062
Balance	<0.001	<0.001	0.0015	0.0029	0.0047
E. coli	0.0013	0.0027	0.0039	0.0051	0.0078
TAE	<0.001	0.0014	0.0028	0.0042	0.0065
Seeds	0.0021	0.0035	0.0048	0.0069	0.0093
CMC	<0.001	<0.001	0.0017	0.0026	0.0043
Hungarian	0.0016	0.0029	0.0045	0.0058	0.0074

As shown in Table 8, the majority of p-values are below the 0.05 significance threshold. This confirms that the performance improvements of the proposed DRL–QWOA–KMeans algorithm are **statistically significant** across the evaluated datasets compared to the baseline methods. Overall, the statistical analysis (mean ± std and Wilcoxon signed-rank test) strongly supports the robustness, stability, and superiority of the proposed hybrid clustering framework over existing metaheuristic-based approaches.

4.7 Runtime and Computational Cost Analysis

To provide a comprehensive evaluation of the proposed method, the computational efficiency of all algorithms was analyzed. All experiments were conducted on the same hardware platform (Intel Core i7-12700H processor, 32 GB RAM, and NVIDIA RTX 3060 GPU) using Python 3.10 and consistent implementation frameworks. Each algorithm was executed 30 times on every dataset under identical conditions (population size = 30, maximum iterations = 100).

Table 9 presents the average runtime (in seconds) along with the standard deviation for each algorithm across all ten UCI benchmark datasets.

Table 9. Average runtime (Mean ± Std) in seconds over 100 independent runs

Dataset	GA-KMeans	PSO-KMeans	DE-KMeans	WOA-KMeans	QWOA-KMeans	DRL–QWOA–KMeans
Hill-Valley	48.7 ± 2.1	41.3 ± 1.8	45.6 ± 2.0	39.8 ± 1.7	42.5 ± 1.9	68.4 ± 3.2
Dermatology	32.4 ± 1.6	28.7 ± 1.4	30.9 ± 1.5	27.5 ± 1.3	29.1 ± 1.4	46.7 ± 2.4
Iris	8.9 ± 0.5	7.6 ± 0.4	8.3 ± 0.5	7.2 ± 0.4	7.8 ± 0.4	12.5 ± 0.8
Wine	11.2 ± 0.6	9.8 ± 0.5	10.7 ± 0.6	9.4 ± 0.5	10.1 ± 0.5	15.9 ± 0.9
Balance	52.6 ± 2.3	44.8 ± 2.0	49.3 ± 2.2	43.1 ± 1.9	45.7 ± 2.0	71.3 ± 3.5
E. coli	19.8 ± 1.1	17.2 ± 0.9	18.6 ± 1.0	16.5 ± 0.8	17.9 ± 0.9	28.4 ± 1.6
TAE	9.7 ± 0.6	8.4 ± 0.5	9.1 ± 0.5	8.1 ± 0.4	8.6 ± 0.5	13.8 ± 0.8
Seeds	13.5 ± 0.7	11.9 ± 0.6	12.8 ± 0.7	11.4 ± 0.6	12.2 ± 0.6	19.1 ± 1.1
CMC	89.4 ± 3.8	76.2 ± 3.2	83.7 ± 3.5	73.9 ± 3.1	78.5 ± 3.3	124.6 ± 5.7
Hungarian	22.6 ± 1.2	19.8 ± 1.0	21.4 ± 1.1	19.1 ± 0.9	20.3 ± 1.0	32.7 ± 1.8
Average	**30.9**	**26.6**	**29.0**	**25.6**	**27.3**	**43.3**

As shown in Table 9, the proposed DRL–QWOA–KMeans algorithm requires higher computational time compared to the baseline metaheuristic methods. On average, it is approximately **1.6 times slower** than the fastest method (WOA-KMeans). This increase in runtime is primarily due to the additional overhead introduced by the Deep Reinforcement Learning agent, which performs forward passes through the neural network and updates the QWOA parameters adaptively at each iteration. However, this moderate computational overhead is well justified by the significant improvements in clustering quality (lower WCSS, higher accuracy, sensitivity, specificity, and F-score) and greater stability (lower standard deviation) achieved by the proposed method. The adaptive parameter control mechanism enables more effective exploration–exploitation balancing, leading to superior solutions that static-parameter algorithms cannot consistently achieve.

5. Conclusion and Future Work

In this study, a novel hybrid clustering framework named DRL–QWOA–KMeans was introduced to overcome the inherent limitations of traditional K-Means such as random initialization sensitivity, premature convergence, and low clustering accuracy. By integrating Deep Reinforcement Learning (DRL) to dynamically balance exploration–exploitation and Quantum Whale Optimization Algorithm (QWOA) to guide centroid updates toward global optima, the proposed method achieved superior performance across multiple UCI benchmark datasets. Quantitative results demonstrated remarkable improvements over classical and hybrid meta-heuristic approaches in terms of WCSS, AAcc, ASen, ASpe, and F-scoreM, while qualitative visual analysis confirmed clearer and more compact cluster boundaries.

Future work will focus on extending the DRL–QWOA–KMeans model to high-dimensional and streaming data environments, developing adaptive reward mechanisms for reinforcement learning, and exploring parallel and quantum-computing implementations to enhance computational scalability and convergence efficiency.

References

1. Jain AK (2008) Data clustering: 50 years beyond k-means. In: *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, pp 3–4.
2. GeethaRamani R, Balasubramanian L (2018) Macula segmentation and fovea localization employing image processing and heuristic-based clustering for automated retinal screening. *Comput Methods Programs Biomed* 160:153–163.
3. Abualigah LM, Khader AT, Al-Betar MA, Alomari OA (2017) Text feature selection with a robust weight scheme and dynamic dimension reduction to text document clustering. *Expert Syst Appl* 84:24–36.
4. Triguero I, del Río S, López V, Bacardit J, Benítez JM, Herrera F (2015) ROSEFW-RF: the winner algorithm for the ECBDL'14 big data competition: an extremely imbalanced big data bioinformatics problem. *Knowl-Based Syst* 87:69–79.
5. Zhu J, Lung C-H, Srivastava V (2015) A hybrid clustering technique using quantitative and qualitative data for wireless sensor networks. *Ad Hoc Netw* 25:38–53.
6. Gan G, Ng MK-P (2017) K-means clustering with outlier removal. *Pattern Recogn Lett* 90:8–14.
7. Raykov YP, Boukouvalas A, Baig F, Little MA (2016) What to do when k-means clustering fails: a simple yet principled alternative algorithm. *PLoS ONE* 11(9):e0162259.
8. Jahwar AF, Abdulazeez AM (2020) Meta-heuristic algorithms for k-means clustering: a review. *PalArch's J Archaeol Egypt/Egyptol* 17(7):12002–12020.
9. Minh H-L, Sang-To T, Wahab MA, Cuong-Le T (2022) A new metaheuristic optimization based on k-means clustering algorithm and its application to structural damage identification. *Knowl-Based Syst* 251:109189.
10. Braik, Malik, et al. "Enhanced whale optimization algorithm-based modeling and simulation analysis for industrial system parameter identification." *Journal of Supercomputing* 79.13 (2023).
11. Bezdan T, Stoean C, Naamany AA, Bacanin N, Rashid TA, Zivkovic M, Venkatachalam K (2021) Hybrid fruit-fly optimization algorithm with k-means for text document clustering. *Mathematics* 9(16):1929.

12. Abasi AK, Khader AT, Al-Betar MA, Naim S, Alyasseri ZAA, Makhadmeh SN (2020) A novel hybrid multi-verse optimizer with k-means for text document clustering. *Neural Comput Appl* 32:17703–17729.
13. Purushothaman R, Rajagopalan S, Dhandapani G (2020) Hybridizing gray wolf optimization with grasshopper optimization algorithm for text feature selection and clustering. *Appl Soft Comput* 96:106651.
14. Braik M, Sheta A, Al-Hiary H (2021) A novel meta-heuristic search algorithm for solving optimization problems: capuchin search algorithm. *Neural Comput Appl* 33:2515–2547.
15. Braik M (2021) A hybrid multi-gene genetic programming with capuchin search algorithm for modeling a nonlinear challenge problem: industrial winding process case study. *Neural Process Lett* 53(4):2873–2916.
16. Kanipriya M, Hemalatha C, Sridevi N, SriVidhya S, Shabu SJ (2022) An improved capuchin search algorithm optimized hybrid CNN–LSTM architecture for malignant lung nodule detection. *Biomed Signal Process Control* 78:103973.
17. Abd Elaziz M, Al-qaness MA, Dahou A, Ibrahim RA, Abd El-Latif AA (2023) Intrusion detection approach for cloud and IoT environments using deep learning and capuchin search algorithm. *Adv Eng Softw* 176:103402.
18. Al-qaness MA, Ewees AA, Fan H, Abualigah L, Elsheikh AH, Abd Elaziz M (2023) Wind power prediction using random vector functional link network with capuchin search algorithm. *Ain Shams Eng J* 14(9):102095.
19. Braik M, Awadallah MA, Al-Betar MA, Hammouri AI, Alzubi OA (2023) Cognitively enhanced versions of capuchin search algorithm for feature selection in medical diagnosis: a COVID-19 case study. *Cognit Comput*, pp 1–38.
20. Li S, Li Z, Li Q, Zhang M, Li L (2023) Hybrid improved capuchin search algorithm for plant image thresholding. *Front Plant Sci* 14:1122788.
21. Rizk-Allah RM, Hassanien AE, Snášel V (2022) A hybrid chameleon swarm algorithm with superiority of feasible solutions for optimal combined heat and power economic dispatch problem. *Energy* 254:124340.
22. Braik MS, Awadallah MA, Al-Betar MA, Hammouri AI, Zitar RA (2023) A non-convex economic load dispatch problem using chameleon swarm algorithm with roulette wheel and Lévy flight methods. *Appl Intell*, pp 1–40.
23. Braik MS, Hammouri AI, Awadallah MA, Al-Betar MA, Khtatneh K (2023) An improved hybrid chameleon swarm algorithm for feature selection in medical diagnosis. *Biomed Signal Process Control* 85:105073.
24. Kumar S, Fred AL, Varghese PS (2019) Suspicious lesion segmentation using optimized spatial feature-based superpixel fuzzy c-means clustering. *J Digit Imaging* 32(2):322–335.
25. Zhao X, Wu Y, Song G, Li Z, Zhang Y, Fan Y (2018) A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. *Med Image Anal* 43:98–111.
26. Dolz J, Desrosiers C, Ayed IB (2018) 3D fully convolutional networks for subcortical segmentation in MRI: a large-scale study. *NeuroImage* 170:456–470.
27. Ali HH, Kadhum LE (2017) K-means clustering algorithm applications in data mining and pattern recognition. *Int J Sci Res* 6(8):1577–1584.
28. Septiarini A, Hamdani H, Sari SU, Hatta HR, Puspitasari N, Hadikurniawati W (2022) Tomato image segmentation using k-means clustering and edge detection. In: *Proc IEEE ISMODE*, pp 92–96.
29. Bouadjenek MR, Sanner S, Du Y (2020) Relevance- and interface-driven clustering for visual information retrieval. *Inf Syst* 94:101592.
30. Likas A, Vlassis N, Verbeek JJ (2003) The global k-means clustering algorithm. *Pattern Recogn* 36(2):451–461.
31. Karaboga D, Ozturk C (2011) A novel clustering approach: artificial bee colony (ABC) algorithm. *Appl Soft Comput* 11(1):652–657.
32. Shelokar P, Jayaraman VK, Kulkarni BD (2004) An ant colony approach for clustering. *Anal Chim Acta* 509(2):187–195.
33. Chen C-Y, Ye F (2012) Particle swarm optimization and its application to clustering analysis. In: *Proc IEEE Conf Electrical Power Distribution*, pp 789–794.
34. Senthilnath J, Omkar SN, Mani V (2011) Clustering using firefly algorithm: performance study. *Swarm Evol Comput* 1(3):164–171.
35. Hatamlou A, Abdullah S, Nezamabadi-Pour H (2012) A combined approach for clustering based on k-means and gravitational search algorithms. *Swarm Evol Comput* 6:47–52.

36. Hatamlou A (2013) Black hole: a new heuristic optimization approach for data clustering. *Inf Sci* 222:175–184.
37. Hatamlou A, Abdullah S, Hatamlou M (2011) Data clustering using big bang–big crunch algorithm. In: *Proc Int Conf Innovative Computing Technology*, Springer, pp 383–388.
38. Das P, Das DK, Dey S (2018) Modified bee colony optimization hybridized with k-means for data clustering. *Appl Soft Comput* 70:590–603.
39. Niknam T, Amiri B (2010) An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis. *Appl Soft Comput* 10(1):183–197.
40. Boushaki SI, Kamel N, Bendjeghaba O (2018) A new quantum chaotic cuckoo search algorithm for data clustering. *Expert Syst Appl* 96:358–372.
41. Zhou L, Li L (2018) Improvement of the firefly-based k-means clustering algorithm. In: *Proc Int Conf Data Science*.
42. Hassanzadeh T, Meybodi MR (2012) A hybrid firefly and k-means clustering approach. In: *Proc IEEE AISP*, pp 7–11.
43. Han X, Quan L, Xiong X, Almeter M, Xiang J, Lan Y (2017) A novel data clustering algorithm based on modified gravitational search algorithm. *Eng Appl Artif Intell* 61:1–7.
44. Nanda SJ, Panda G (2014) A survey on nature-inspired metaheuristic algorithms for partitional clustering. *Swarm Evol Comput* 16:1–18.
45. Gupta A, Datta S, Das S (2019) Evolutionary multiobjective fuzzy clustering at different fuzziness levels. *IEEE Trans Cybern* 51(5):2601–2611.
46. Lai Z, Feng X, Yu H, Luo F (2018) Parallel social spider optimization algorithm based on emotional learning. *IEEE Trans Syst Man Cybern Syst* 51(2):797–808.
47. Yan D, Cao H, Yu Y, Wang Y, Yu X (2020) Cat swarm optimization-based clustering for single- and multiobjective data partitioning. *IEEE Trans Autom Sci Eng* 17(3):1633–1646.
48. Blake C, Merz C (2015) Uci repository of machine learning databases, department of information and computer science, university of california, irvine, ca, URL: http://www. archive. ics. uci. edu/ml