



A SYSTEMATIC REVIEW OF MACHINE LEARNING-ENHANCED METAHEURISTICS FOR SOLVING CAPACITATED VEHICLE ROUTING PROBLEMS

Revisión sistemática de metaheurísticas mejoradas con aprendizaje automático para resolver problemas de enrutamiento de vehículos con capacidad

Mauricio Maca-Chagüendo 
Universidad del Cauca (Popayán, Colombia).
mmaca@unicauca.edu.co

Carlos-Alberto Cobos-Lozada 
Universidad del Cauca (Popayán, Colombia).
ccobos@unicauca.edu.co

Fecha de recibido: 10-03-2024

Fecha de aceptado: 21-06-2024



ABSTRACT

Combinatorial optimization problems are prevalent in various domains, such as logistics, planning, and resource allocation. This study is essential to fields such as applied mathematics and computer science. A prominent example is the Capacitated Vehicle Routing Problem (CVRP), which aims to minimize the route costs for a fleet of vehicles with capacity constraints. Traditionally, this problem has been tackled using exact methods and metaheuristics, but advancements in hybrid algorithms, modern metaheuristics, and machine learning have significantly improved both the solution quality and computational efficiency. This review followed Kitchenham's methodological guidelines, selecting and analyzing articles published between 2019 and 2024 from the Scopus and Web of Science databases. Specific inclusion and quality assessment criteria were applied, synthesizing the information quantitatively and qualitatively and categorizing studies based on their approaches and techniques. The studies were grouped by their focus, addressing research questions on advancements in machine-learning techniques for enhancing metaheuristics, datasets used, and comparison metrics. The findings emphasize the potential of hybrid approaches that combine metaheuristics with machine learning techniques to effectively solve the CVRP. These approaches have demonstrated their ability to improve efficiency and solution quality, which is particularly relevant for large-scale, complex problems. Significant improvements in solution accuracy and considerable reductions in computation times were observed.

Keywords: artificial neural networks; Capacitated Vehicle Routing Problem (CVRP); machine learning; metaheuristics; reinforcement learning.

RESUMEN

Los problemas de optimización combinatoria son muy comunes en diversas áreas como la logística, la planificación y la asignación de recursos. Este estudio es esencial en campos como las matemáticas aplicadas y las ciencias de la computación. Un ejemplo destacado es el problema de enrutamiento de vehículos con capacidad (CVRP), que busca minimizar los costos de las rutas de una flota de vehículos con restricciones de capacidad. Tradicionalmente, este problema se ha abordado mediante métodos exactos y metaheurísticos, pero la evolución de algoritmos híbridos, las metaheurísticas modernas y el aprendizaje

automático han mejorado tanto la calidad de las soluciones como la eficiencia computacional. La revisión se realizó siguiendo las directrices metodológicas de Kitchenham. Se seleccionaron y analizaron artículos publicados entre 2019 y 2024 en bases de datos como Scopus y Web of Science. Se aplicaron criterios específicos de inclusión y evaluación de la calidad, sintetizando la información de forma cuantitativa y cualitativa, y categorizando los estudios según sus enfoques y técnicas empleadas. Los estudios se agruparon según su enfoque y se abordaron las preguntas de investigación sobre avances en técnicas de aprendizaje automático para mejorar las metaheurísticas, los conjuntos de datos empleados y las métricas de comparación. Los resultados destacan el potencial de los enfoques híbridos que combinan metaheurísticas con técnicas de aprendizaje automático para resolver el CVRP de manera efectiva. Estos enfoques han demostrado ser capaces de mejorar la eficiencia y la calidad de las soluciones, lo cual es especialmente relevante en problemas de gran escala y complejidad. Se observó una mejora significativa en la precisión de las soluciones y una reducción considerable en los tiempos de cómputo.

Palabras clave: aprendizaje automático; aprendizaje por refuerzo; metaheurísticas; problema de enrutamiento de vehículos con capacidad; redes neuronales artificiales.

1. INTRODUCTION

Combinatorial Optimization Problems (COPs) are foundational across disciplines such as applied mathematics, computer science, operations research, engineering, and economics. They seek to maximize or minimize an objective function over a finite but vast set of solutions, which makes them computationally complex, with many classified as NP-hard [1]-[2]. Theoretically, COPs have been studied in complexity theory and areas of mathematics such as graph theory, probability, and optimization, which has enabled the formulation, analysis, and development of methods to find approximate solutions in polynomial time [3]-[4].

In practice, COPs are applied to route optimization, resource allocation, scheduling, and inventory management, which are crucial for improving processes. In this context, the Capacitated Vehicle Routing Problem (CVRP) stands out as one of the most researched and applied problems, as it optimizes distribution routes with capacity constraints and minimizes travel costs [2]-[4]. Exact and heuristic methods have traditionally been used to address this problem. However, exact methods, such as integer linear programming and branch-and-bound techniques, are only suitable for small instances. By contrast, heuristic and metaheuristic methods, such as the Clarke-Wright algorithm, tabu search, and variable neighborhood search, are preferred for large-scale problems because they can approximate optimal solutions at a lower computational cost [5]-[6].

The CVRP solutions have recently evolved using advanced artificial intelligence (AI) and machine learning (ML) techniques. Modern metaheuristics, such as Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO), along with hybrid and memetic techniques, have improved the efficiency of exploring the solution space. Additionally, the integration of neural networks and reinforcement learning (RL) has led to neural solvers capable of learning patterns and strategies, enabling a more efficient resolution of complex CVRP instances in terms of time and solution quality [7].

Thirty articles published between 2019 and 2024 were selected from the Web of Science and Scopus scientific databases for this study. The documents were categorized by the type of metaheuristics employed, machine-learning technique applied, and identification of datasets used for evaluation.

The rest of this paper is organized as follows: Section two summarizes the methodology followed, including the research question, search strategy, selection criteria for primary studies, quality evaluation criteria, data extraction, and synthesis method. Section three presents the results related to metaheuristics and the various machine learning techniques. Finally, section four presents the conclusions of the systematic review.

2. METHODOLOGY

This systematic review was conducted following the methodological guidelines established in Kitchenham and Brereton [8], given their rigor and widespread use in various fields of knowledge. This methodology consists of three main phases: planning, execution of the review, and dissemination of results.

A. Research Question

This work aims to select, analyze, and synthesize documents that address approaches, techniques, and models for solving the CVRP, determining recent advancements that enhance metaheuristics with machine learning techniques. [Table 1](#) shows the research questions formulated to guide the collection and categorization of information, identifying the most promising and notable proposals to date.

Table 1. Research questions

Research Questions	Motivation of the Question
Q1. What machine learning techniques are used, how do they operate, and what stage do they support in a metaheuristic algorithm to solve the CVRP problems?	This question aims to identify and understand the specific machine learning techniques used to enhance metaheuristic algorithms for the CVRP and their role within these algorithms.
Q2. How does the metaheuristic algorithm use the knowledge obtained from the machine learning technique to solve the CVRP problems?	This question aims to identify how metaheuristic algorithms integrate and apply the knowledge generated by machine learning techniques to improve the CVRP resolution.
Q3. What CVRP datasets are used to evaluate and compare the performance of the algorithms?	The aim is to identify the datasets/problems used in the reviewed studies, which is crucial for understanding the validity and comparability of results across different approaches.
Q4. What metrics are used to evaluate and compare the performance of the algorithms?	This question seeks to identify the specific metrics used to evaluate and compare the effectiveness and efficiency of algorithms enhanced by machine learning techniques in solving the CVRP.

B. Search Strategy

The article search was conducted using Scopus and Web of Science. [Table 2](#) details the search string, which was iteratively refined after analyzing the preliminary results. Based on the PICOC method (Population, Intervention, Comparison, Outcomes, Context) [9], the search string was structured with terms combined by conjunction and applied to titles, abstracts, and keywords, using advanced search options to optimize the results.

C. Selection Criteria for Primary Studies

Specific inclusion criteria were applied, focusing on the titles, abstracts, and keywords. The selected studies had to meet the following inclusion criteria: 1) be journal articles or conference papers, 2) be written in English or Spanish, 3) have been published between 2019 and 2024, 4) address at least one of the formulated research questions, and 5) be freely accessible or available through the University of Cauca's library. The selected studies were evaluated after removing duplicate documents and applying these selection criteria.

Table 2. Advanced Search Strings

Source	Search String	Articles
Scopus	TITLE-ABS-KEY = ((metaheuristic OR heuristic OR genetic OR swarm OR “optimization algorithm” OR “Evolutionary computation” OR “memetic algorithm” OR “differential evolution” OR “hybrid metaheuristic” OR “Combinatorial optimization”) AND (“capacitated vehicle routing problem” OR CVRP) AND (“aprendizaje de máquina” OR “artificial intelligence” OR “deep learning” OR “learnable evolution model” OR LEM OR “data mining” OR “reinforcement learning”)) AND PUBYEAR > 2018 AND (LANGUAGE (english) OR LANGUAGE (spanish)) AND DOCTYPE (“cp” OR “ar”)	31
Web of Science	TS = ((metaheuristic OR heuristic OR genetic OR swarm OR “optimization algorithm” OR “Evolutionary computation” OR “memetic algorithm” OR “differential evolution” OR “hybrid metaheuristic” OR “Combinatorial optimization”) AND (“capacitated vehicle routing problem” OR CVRP) AND (“aprendizaje de máquina” OR “artificial intelligence” OR “deep learning” OR “learnable evolution model” OR LEM OR “data mining” OR “reinforcement learning”)) AND PY=(2019-2024)	22

D. Quality Evaluation Criteria

Table 3 details the six quality criteria established to evaluate the primary studies using a three-point scoring system (0, 0.5, 1). In this system, a score of zero (0) indicates that the article does not meet the quality criterion, a score of zero point five (0.5) indicates partial fulfillment of the criterion, and a score of one (1) indicates full compliance. The studies received a minimum score of 0 and maximum score of 6. The evaluation was conducted using the full texts, and the answers to the research questions allowed for classifying the studies according to their approach and content, thus providing an overview of the state of knowledge in the field.

Table 3. Quality Criteria

Criterion	Description	Score		
		0	0.5	1
C1	The study integrates metaheuristic approaches with machine learning techniques.	No	Partially	Yes
C2	Other authors have cited the article.	No	Partially	Yes
C3	The article describes the algorithms and their interaction to solve the problem in detail.	No	Partially	Yes
C4	The experimental design for evaluation and comparison allows the experiment to be replicated.	No	Partially	Yes
C5	The article presents in detail the results of the experimentation.	No	Partially	Yes
C6	The article presents coherence between the proposal, the experimentation, the discussion, and the conclusions.	No	Partially	Yes

E. Synthesis Methods and Systematic Review Timeline

The synthesis was conducted using both quantitative and qualitative data. The studies were organized into a table by year, reference, approach, techniques, repositories, and quality criteria. The studies were grouped into three main categories, and within each, a summary of the central primary studies found is presented by year. Data collection for this study was conducted between May and July 2024.

3. RESULTS AND DISCUSSION

A total of 53 documents were found through Web of Science and Scopus. After applying the selection criteria, 32 relevant studies were identified after applying selection criteria. Subsequently, two articles were discarded after the quality evaluation, leaving 30 studies, which were classified into three categories: a) results related to metaheuristics (see Table 4), b) results related to reinforcement learning (see Table 5), and c) results related to machine learning (see Table 6). The following subsections describe each group of studies and highlight the main elements and quality evaluation results.

A. Results Related to Metaheuristics

The CVRP-FA algorithm was introduced in 2019 as an improved variant of the Firefly Algorithm (FA) for the CVRP. This algorithm enhances the convergence and solution quality by integrating the local search (2-Opt and 2-h-opt) and genetic operators. Evaluated on 82 instances from the CVRPLIB (<http://vrp.galagos.inf.puc-rio.br/index.php/en/>), the CVRP-FA demonstrated faster convergence and greater precision than previous FA versions, achieving its effectiveness through parameter tuning using the Taguchi method [1].

In 2019, an improved algorithm based on the Fireworks Algorithm (FWA) was introduced to solve the CVRP. The FWA generates new solutions through explosions and mutations, thereby improving route quality. Evaluated on the P and G instances from the CVRPLIB, it was compared with algorithms like Active-Guided Evolution Strategies (AGES), Greedy Randomized Evolutionary Local Search (GRELS), and Hybrid Genetic Particle Swarm Optimization (HGPSO), demonstrating superior solution quality with minimal deviation from benchmark solutions and faster execution times [10].

In 2020, researchers introduced a hybrid algorithm that integrates ACO with Variable Neighborhood Search (VNS) to address the CVRP in scenarios focused on equitable distribution route scheduling for emergency aid. The approach generates initial solutions with ACO and refines them with VNS, optimizing the waiting times. Compared to ACO with 2-opt, GA, and CPLEX, it improved equity and solution quality. The test data randomly generated based on previous studies allowed the evaluation of its effectiveness in various scenarios [11].

In 2021, the hybrid algorithm MAS-SA-PR was proposed, integrating a Modified Ant System (MAS), Sweep Algorithm (SA), and Path Relinking (PathR) to solve the CVRP. The algorithm converts customer locations to polar coordinates, initializes routes, and uses local search techniques to improve solutions. The MAS-SA-PR demonstrated competitive results on A, B, and P instances from the CVRPLIB compared to the Large Neighborhood Search–Ant Colony Optimization (LNS-ACO), Ranking-Based Monte Carlo–Artificial Bee Colony (RBMC-ABC), and the CVRP-FA algorithms. However, the lack of a formal statistical analysis limits the validity of its conclusions [4].

In 2021, the authors continued their previous work [1] by proposing a collaborative approach to solve the CVRP called the Cooperative Firefly Algorithm (CHFA). First, they designed a new hybrid FA (HFA) by modifying the local search and mutation from the CVRP-FA. Then, they created a parallel cooperative model with multiple HFAs, where firefly populations periodically share solutions to improve convergence and avoid stagnation. Evaluated on 108 instances from the CVRPLIB, the algorithm demonstrated improvements in solution quality and computational speed compared to the CVRP-FA, LNS-ACO, HA, and Discrete Invasive Weed Optimization (DIWO) algorithms, and performed comparably to the Improved Symbiotic Organism Search (ISOS) and Gravitational Emulation Local Search (CVRP_GELS) algorithms [2].

In 2022, BACO (Bilevel ACO), an algorithm based on ACO, was introduced to solve the Capacitated Electric Vehicle Routing Problem (CEVRP). The algorithm operates on two levels: the upper level generates solutions for the CVRP subproblem, and the lower level, with routes that meet capacity constraints, solves the Fixed Route Vehicle Charging Problem (FRVCP). Evaluated in the IEEE WCCI2020 competition, BACO updated the best solutions in seven out of ten large instances and showed fast convergence compared to algorithms like Iterative Local Search (ILS) and a single-level Max-Min Ant System (SLMMAS) [12].

In the same year (2022), ADACO was introduced, which integrates Adaptive Stochastic Gradient Descent (ASGD) into ACO to improve the CVRP solution quality. ADACO dynamically updates the pheromone values, improving convergence and avoiding premature convergence. Evaluated on instances from the TSPLIB and CVRPLIB, the algorithm outperformed the Ant Colony System (ACS), the Max-Min Ant System (MMAS), and the ILS regarding solution precision and convergence. The Wilcoxon test confirmed that the results were statistically significant in most cases [5].

In 2022, a Neural Large Neighborhood Search (NLNS) was introduced, combining a Large Neighborhood Search (LNS) with deep neural networks and Attention Mechanisms (AM) to repair incomplete CVRP solutions. Evaluated on various instances, it performed similarly to LKH3, with lower-cost solutions in several cases but did not surpass UHGS. For the SDVRP, SplitILS outperformed NLNS on smaller instances but outperformed it on larger ones. The results of the Capacitated Team Orienteering Problem (CTOP) were comparable to those achieved by the most advanced approaches. The study was initially presented in [10] and expanded in [7].

Additionally, in 2022, an improved variant of ACO called DSRACO was introduced to solve the CVRP, which incorporates methods like Dynamic Space Reduction (DSR), an elitist mechanism, and large-scale neighborhood search operators. This approach improves the solution quality and prevents the algorithm from becoming stuck in the local optima. The evaluation of 73 CVRPLIB instances showed that DSRACO improves the solution efficiency and quality compared to ACO and competes with LNS-ACO, CVRP-FA, and the Order-aware Hybrid Genetic Algorithm (OHGA) [3].

In 2023, the General Variable Neighborhood Search (GVNS) metaheuristic was enhanced by integrating Multi-Armed Bandit (MAB) algorithms and variations of the Upper Confidence Bound (UCB) algorithm to optimize neighborhood selection, creating a scheme called Bandit VNS. Evaluated on CVRPLIB sets, Bandit VNS improved both solution quality and execution speed, demonstrating that combining reinforcement learning and metaheuristics is a promising strategy for solving routing problems [6].

In 2023, a hybrid algorithm combining clustering techniques with ACO was introduced to solve the Energy Minimizing Vehicle Routing Problem (EMVRP). K-means and M-medoids were applied to improve ACO efficiency by reducing the complexity of the problem. Four hybrid algorithms were designed: Free Ant + K-Means and Free Ant + K-Medoids, which explored multiple clusters, and Restricted Ant + K-Means and Restricted Ant + K-Medoids, limited to routing within preformed clusters. Evaluated on CVRPLIB instances, Free Ant + K-Means excelled in solution quality. Nevertheless, they required more computation time, while the Restricted Ant algorithms were more efficient and suited for scenarios with strict time limits [13].

Among the results related to metaheuristics for solving the CVRP, it was found that some have been improved by integrating machine learning techniques. For example, the LNS metaheuristic was improved by guiding the repair of incomplete solutions by combining a Deep Neural Network (DNN) and AM techniques [7]. The ASGD was integrated into the ACO metaheuristic to improve solution

quality [5]. This metaheuristic was also enhanced by integrating clustering techniques like M-Means and K-Medoids to improve efficiency [13]. Finally, the GVNS metaheuristic was strengthened using RL, applying the MAB approach and UCB implementation to optimize neighborhood selection [6].

Table 4. Results Related to Metaheuristics

Year	Reference	Techniques	Repositories	Quality Criteria						Score
				C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	
2019	[1]	FA	CVRPLIB	0	1	1	0.5	1	1	4.5
2019	[10]	FWA	CVRPLIB	0	1	1	0.5	1	1	4.5
2020	[11]	ACO, VNS	Generated	0	1	1	0.5	1	1	4.5
2021	[4]	MAS, SA, PathR	CVRPLIB	0	1	1	0.5	1	1	4.5
2021	[2]	FA, Parallel	CVRPLIB	0	1	1	0.5	1	1	4.5
2022	[12]	BACO, OS-MMAS, RH	WCCI2020 (IEEE)	0	1	1	0.5	1	1	4.5
2022	[5]	ML, ACO, ASGD	CVRPLIB TSPLIB CVRP	1	1	1	0.5	1	1	5.5
2022	[7]	ML, LNS, DNNs, AM SplitILS	SDVRP CTOP	1	0.5	1	0.5	1	1	5.0
2022	[3]	ACO, DSR, LNS	CVRP	0	1	1	0.5	1	1	4.5
2023	[6]	RL, GVNS, MAB, UHGS	CVRPLIB	1	1	1	0.5	1	1	5.5
2023	[13]	ML, K-Mean, K-Medoids, ACO	CVRPLIB	1	1	1	0.5	1	1	5.5

B. Results Related to Reinforcement Learning

In 2021, a method based on Deep Reinforcement Learning (DRL) was introduced to solve TSP, CVRP, and Multiple Routing with Fixed Fleet Problems (MRPFF). The ERRL technique (Entropy Regularised Reinforcement Learning) includes an entropy term in the policy loss function to encourage exploration and avoid convergence to suboptimal policies. Evaluated using Nazari's instances for the TSP and the CVRP and generated instances for MRPFF, ERRL showed improvements in solution quality and accelerated optimization compared to other approaches [14].

In 2022, a deep reinforcement learning architecture was combined with an attention mechanism to solve the Heterogeneous CVRP (HCVRP). The network policy includes a vehicle selection decoder, a node selection decoder, and an encoder to process the problem's features. The authors presented the MM-HCVRP and MS-HCVRP formulations for the min-max and min-sum objectives. For the experimentation, two fleets were generated: the first with three vehicles with capacities of 20, 25, and 30, and the second with five vehicles with capacities of 20, 25, 30, 35, and 40. The software is available at https://github.com/jingwenli0312/HCVRP_DRL. The results were competitive compared to traditional methods [15].

In 2022, a proposal focused on learning the solution improvement process instead of the solution construction. It uses an RL-based approach to learn policies that guide local search by employing a neural architecture with Self-Attention Mechanisms (Self AM) to parameterize the learning policy. Randomly generated instances of the CVRP and TSP were used, and the evaluation showed competitive results compared to state-of-the-art methods. The advantage of this approach is that it automatically designs a policy, eliminating the need for manual design, which requires substantial domain knowledge [16].

In the same year (2022), DRL with a Multiple Relational Attention Mechanism (MRAM) was applied to improve CVRP solutions. This framework uses relational attention mechanisms to capture the state transition dynamics and enhance node selection in route construction. The model manages state representations by adding relational structures used in decision-making to optimize routes. Evaluations on randomly generated instances for the TSP, CVRP, and SDVRP showed improved solution quality and computational efficiency compared with algorithms such as Gurobi, LKH, and the attention model AM [17].

Continuing in 2022, a Residual Edge-Graph Attention Neural Network (Residual E-GAT) was applied within a DRL framework to improve CVRP solutions. This approach builds on a Graph Attention Network (GAT), considering edge information and residual connections between layers, and optimizing route construction through a decoder that selects nodes based on model-generated probabilities. Experiments were conducted using up to 100 randomly generated nodes from the TSPLIB and CVRPLIB libraries. E-GAT's results surpassed algorithms like PtrNet, GAT, and the attention model AM in terms of optimality and computational efficiency, excelling in generalization and large-scale problem solving [18].

In 2022, a Q-learning-based approach was introduced to solve the School Bus Routing Problem (SBRP), a variant of the CVRP. This approach uses a reinforcement learning model to create the HHQL algorithm, which dynamically selects the best low-level heuristics, such as one-point move, two-point swap, 2-opt, cross, or-opt, two-edge swap, and ruin-and-recreate, based on the accumulated rewards. The algorithm was compared to other heuristic-based approaches, showing that the combination of reinforcement learning and hyper-heuristic selection offers better results in terms of the total distance traveled and the number of routes [19].

In 2023, a General Search Framework (GSF) was introduced, automating the design of search algorithms using reinforcement learning techniques such as Q-learning, Deep Q-Network (DQN), and Proximal Policy Optimization (PPO) to solve the CVRP with Time Windows (CVRPTW). The Actor-Critic with Entropy (ACE) method was introduced to balance exploration and exploitation in learning. The research focused on learning evolutionary operators, heuristic evolution, and the simultaneous learning of these approaches. Instances from the Solomon dataset were evaluated using three ACE models: ACE_FS, ACE_NLAS, and ACE_LAS, each with a different entropy coefficient adjustment scheme. The results showed that GSF with maximum entropy is effective in algorithm design and generalizes well to CVRPTW [20].

In 2023, combining reinforcement learning with genetic algorithms (GA) showed promise for combinatorial optimization problems. The NeuroCrossover algorithm, based on RL, PPO, and the Cross Information Synergistic Attention (CISA) model, intelligently selects the genetic locus for configuring the crossover operator parameters in GA. The research was used to order two-point crossover operators on instances of the CVRP, the TSP, and the Bin Packing Problem (BPP) from the TSPLIB and CVRPLIB libraries. Experimental results showed that NeuroCrossover improves solution quality, convergence rate, and generalization [21].

In addition, in 2023, a hybrid model based on Deep Reinforcement Learning (DRL) was introduced, integrating an attention-based encoder and an LSTM decoder to optimize route coordination in the Traveling Salesman Problem with Drone (TSP-D). It was assumed that the drone is twice as fast as the truck, and the hybrid model's performance was compared with that of the AM model and the TSP-ep-all. The evaluation used datasets with randomly generated and real-world locations from Amsterdam. The results showed improvements in the solution quality and convergence speed, confirming the effectiveness of the hybrid model compared to the reference methods [22].

In the same year (2023), RL based on Proximal Policy Optimization (PPO) was used to automate the design of algorithms for solving the CVRPTW. This approach introduces two groups of features: search-dependent and instance-dependent, optimizing relevant search space information for algorithm design. To solve problem instances, the method uses RL to automatically generate search algorithms in five stages, three related to states, actions, and reward schemes. The evaluation with instances from the Solomon library, using features such as vehicle capacity, time window density, and customer distribution, showed improvements in the number of vehicles and total distance traveled compared to traditional methods [23].

Another 2023 study focused on the automatic design of metaheuristics using RL to solve the CVRPTW. This paper proposes the first General Search Framework (GSF) with a general search structure that allows the formulation and analysis of metaheuristics, facilitating the automatic composition of algorithms using DQN networks and the PPO policy optimization method. These techniques select effective combinations of evolutionary operators at distinct stages of the optimization process to improve metaheuristic performance. The evaluation conducted on CVRPTW instances from the Solomon library shows that the proposed model outperforms manually designed algorithms, highlighting the effectiveness and generalization of the RL-based approach [24].

In 2024, the Deep Dynamic Transformer Model (DDTM) was introduced to solve the Multi-start Team Orienteering Problem (MSTOP) in mission replanning for Unmanned Aerial Systems (UAS). It uses deep reinforcement learning with a modified version of the REINFORCE algorithm to improve learning efficiency. It employs a self-attention policy network for each part of the route, with attention paid between the partial route and the remaining nodes. The experiments demonstrate that DDTM outperforms methods such as the Policy Optimization with Multiple Optima (POMO), offering better results on classic problems like the CVRP and the TSP, and demonstrates its potential in complex mission replanning scenarios [25].

Table 5. Results related to Reinforcement Learning

Year	Reference	Techniques	Repositories	Quality Criteria						Score
				C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	
2021	[14]	DRL, ERRL	Nazari TSP-CVRP MRPF	0	0.5	1	0.5	1	1	4
2022	[15]	DRL AM	Generated CVRPLIB	0	1	1	1	1	1	5.0
2022	[16]	DRL AM	Generated	0.5	1	1	0.5	1	1	5.0
2022	[17]	DRL, MA	Generated	0	1	1	0.5	1	1	4.5
2022	[18]	DRL GAT GANN	Generated TSPLIB, CVRPLIB	0	1	1	0.5	1	1	4.5
2022	[19]	QL	Real instances SBRP	0.5	0.5	1	0.5	1	1	4.5
2023	[20]	GSF, QL, DQL, PPO	SOLOMON	0.5	1	1	0.5	0.5	1	4.5
2023	[21]	CISA, PPO, GA, MDP	TSPLIB, CVRPLIB, BPP	0.5	1	1	0.5	1	1	5.0
2023	[22]	MA, LSTM	Generated Amsterdam	0	1	1	0.5	1	1	4.5
2023	[23]	PPO	SOLOMON	0.5	1	1	0.5	1	1	5.0
2023	[24]	PPO, DQN	SOLOMON	0.5	1	1	0.5	1	1	5.0
2024	[25]	DRL, MSTOP	CVRLIB TSP	0	1	1	0.5	1	1	4.5

C. Results Related to Machine Learning

In 2021, a branch-and-price framework based on ML was introduced to solve VRP variants, such as SVRPWT (Sampled VRPTW) and SCVRP (Sampled CVRP), where customers are a random subset. Random forest and regression forests were used to predict variables and branching scores, integrating node and variable selection policies into the branch-and-price algorithm. Salomon and Gehring & Homberger data, with customers randomly distributed, clustered, or mixed, were evaluated. The results showed reduced search tree sizes and shorter execution times compared to traditional algorithms [27].

In 2022, transfer learning (TL) was applied to neural combinatorial optimization to solve the CVRP. Using a model pre-trained for TSP, TL accelerated learning and improved the results for the CVRP. The model was trained on TSP and then adapted to the CVRP, enhancing the efficiency and quality of the solutions. Experiments showed that models using TL converged faster and achieved shorter routes than those trained from scratch [28].

In 2023, ML was combined with optimization algorithms to solve the CVRP problem, using K-means to group customers, and then applying the nearest neighbor heuristic (NN) and branch-and-cut technique (B&C). K-means clusters customers based on location and vehicle capacity, reducing complexity and improving routing efficiency. The NN and B&C then determine the best route within each cluster. The evaluation of artificially generated instances showed that K-means combined with B&C is more efficient than NN regarding total distance and computation time [29].

In 2023, an approach was proposed to solve the CVRP using a Graph Convolutional Network (GCN) and beam search, formulated as Quadratic Unconstrained Binary Optimization (QUBO) assisted by quantum computing. Randomly generated instances based on Solomon's data were used. The evaluation showed that this approach achieved results similar to LKH3 and Google OR-Tools for small and medium-sized instances, but was outperformed on larger instances [30].

Another 2023 proposal introduced a Real-time Algorithm Configuration (RAC) approach using a gray-box model that combines cost-sensitive machine learning with a dynamic process to eliminate low-performing configurations. This improves efficiency by freeing up resources and adjusting the configuration during execution. Experiments with the Hybrid Genetic Search (HGS) algorithm for solving CVRP using XE CVRP data generated from Vidal and Uchoa instances showed that this strategy allows HGS to reach high-quality solutions in less time [31].

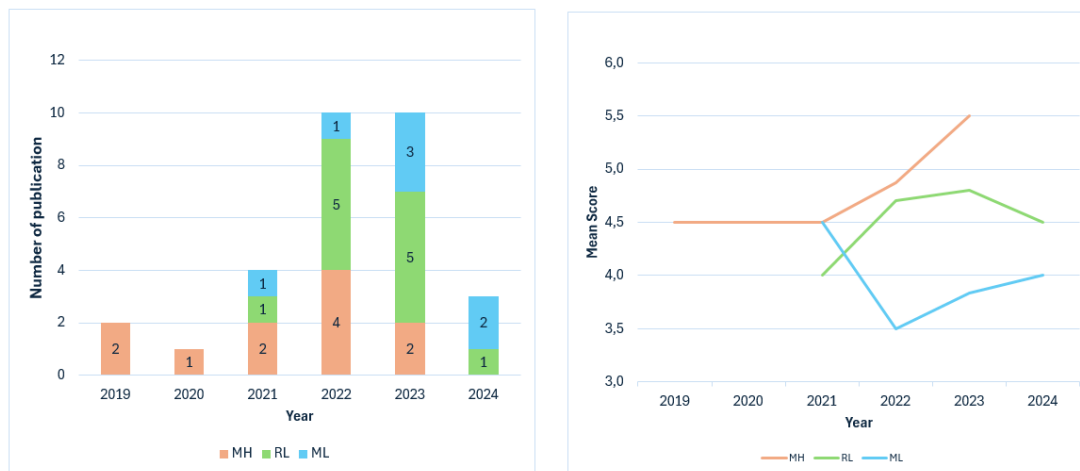
In 2024, the ASP method was introduced, integrating two main components: Distributional Exploration (DE) and Persistent Scale Adaptation (PSA). DE uses a Policy Space Response Oracle (PSRO) and PSA employs a stepwise adaptive procedure. The ASP was evaluated by comparing solvers with and without ASP. The results for the TSPLIB and CVRPLIB instances showed significant improvements in optimality gaps and computational efficiency. Although focused on TSP and CVRP, ASP showed promising results for other combinatorial optimization problems [32].

In 2024, a heuristic approach based on machine learning was introduced to solve instances of the capacitated vehicle routing problem in scenarios with slight changes in customer demands. The algorithm uses a Supervised Learning Model (SLM), specifically, a binary classification model, to predict which edges from a previous solution are likely to remain after demand changes. Compared to the FILO algorithm (Fast Iterated Localized Search Optimization), the machine learning and heuristic approach allowed for high-quality solutions in less time. Evaluations conducted with instances from the X set of the CVRPLIB and modified configurations showed significant reductions in re-optimization times, with optimality gaps ranging from 0% to 1.7% [33].

Table 6. Results Related to Machine Learning

Year	Reference	Techniques	Repositories	Quality Criteria						Score
				C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	
2021	[27]	Random forest classifiers, Regression Forest Estimator	Salomon Gehring & Homberger	0	1	1	0.5	1	1	4.5
2022	[28]	ANN, NCO, AM, TL	Generated	0	0.5	1	0.5	0.5	1	3.5
2023	[29]	K-Means, NN B&C	Generated	0	0	1	0.5	1	1	3.5
2023	[30]	Beam Search, GCN	Generated SOLOMON	0	1	1	0.5	1	1	4.5
2023	[31]	RAC, Gray-Box	Generated XE CVRP	0	0	1	0.5	1	1	3.5
2024	[32]	RL, ASP, PSA, DE	TSPLIB, CVRPLIB, SDVRP, Generated. Real world	0	0.5	1	0.5	1	1	4.0
2024	[33]	SLM	CVRPLIB	0	0.5	1	0.5	1	1	4.0

Among the 30 articles reviewed, 40% (12 studies) focused on reinforcement learning, 37% (11 studies) on metaheuristics, and 23% (7 studies) on machine learning. [Figure 1a](#) shows the distribution of analyzed publications by year organized into these three categories, where an increase in the use of reinforcement learning can be observed. Additionally, as shown in [Figure 1b](#), metaheuristics achieved, on average, a higher score than reinforcement learning, followed by machine learning, according to the evaluation criteria employed.

**Figure 1.** (a) Distribution of Publications by Year. (b) Mean Score of Publications by Year.

D. Response to Research Questions

The research questions are answered below based on the previously presented description of the selected and analyzed primary studies.

1) What machine learning techniques are being used, how do they operate, and which stage of a metaheuristic algorithm do they support in solving the CVRP problems? The main ML techniques used to enhance metaheuristics in solving CVRP problems include DNNs, MA, ASGD, clustering, and RL with MAB and UCB approaches [4]-[7], [13].

2) How does the metaheuristic algorithm uses the knowledge obtained from the machine learning technique to solve the CVRP problems? The knowledge obtained from ML techniques is used to guide the search, improve neighborhood selection, repair incomplete solutions, enhance pheromone values to prevent premature convergence and optimize adaptive neighborhood selection to enable more efficient searches focused on promising areas [5]-[7].

3) What CVRP datasets are used to evaluate and compare the performance of the algorithms? The most used datasets to evaluate and compare algorithm performance are those provided in CVRPLIB and Solomon [30], [33]. Randomly generated instance sets from each study are also employed [16], [18].

4) What metrics are used for the evaluation and comparison of the algorithm performance? The percentage deviation (Gap) is used, including the Best Gap (for the best solution obtained) and the Average Gap, to measure the average quality of the solutions found. To evaluate computational efficiency, execution time, number of iterations, and number of objective function evaluations are used [5], [7].

4. CONCLUSIONS AND FUTURE WORK

The techniques, methods, and models used to solve the CVRP problem through metaheuristic algorithms, reinforcement learning, and machine learning were identified and synthesized in this systematic review. A notable trend in the development of hybrid computational models has emerged, specifically those that integrate machine-learning techniques with metaheuristics to improve solution quality and computational efficiency. In the metaheuristic-related studies analyzed, four out of nine employed this hybrid approach, and their results indicate that this integration is promising. These models combine the exploration and adaptation capabilities of metaheuristics with the learning and prediction abilities of machine-learning techniques, allowing the search process to be more efficient and precise in complex problems, such as CVRP.

Reinforcement learning-based models for solving the CVRP have gained attention in recent years. By integrating deep learning with attention mechanisms and neural networks, processes such as node selection and route construction have been improved, outperforming traditional algorithms. Reinforcement learning has been pivotal in automating the design of metaheuristics, as seen in GSF, which generates optimized algorithms for the CVRP with better generalization and effectiveness. This has enabled more efficient solution space exploration, avoiding premature convergence to suboptimal policies and improving both the quality of solutions and computational efficiency.

In other machine learning studies, techniques such as Random Forests, TL, GCN, K-Means, and Branch & Cut, combined with optimization algorithms, have effectively reduced search tree size and accelerated learning, highlighting the ML's potential in solving routing problems. Most experiments used standard datasets like TSPLIB, CVRPLIB, and Solomon, with solution quality evaluated using percentage deviation and computational efficiency measured by execution time, iteration count, and objective function evaluations.

Based on the findings of this systematic review, future research should prioritize the development of an integrated framework that enhances existing metaheuristics through the integration of machine-

learning techniques. This framework would systematically identify areas within the metaheuristic that could benefit from ML optimization, such as neighborhood selection, solution repair, or algorithm convergence.

AUTHORS' CONTRIBUTION

Mauricio Maca-Chagüendo: Methodology; Investigation; Formal analysis; Writing – original draft; Writing – review & editing.

Carlos-Alberto Cobos-Lozada: Conceptualization; Formal analysis; Supervision; Writing – review & editing.

FUNDING

This scientific research article is derived from the research project “Machine Learning as an Enhancement to Metaheuristic Algorithms for Solving Capacitated Vehicle Routing Problems,” partially funded by the University of Cauca”. The project started in 2023 and will be completed in 2027.

REFERENCES

- [1] A. M. Altabeeb, A. M. Mohsen, A. Ghallab, “An improved hybrid firefly algorithm for capacitated vehicle routing problem,” *Applied Software Computing*, vol. 84, e105728, Nov. 2019.
<https://doi.org/10.1016/j.asoc.2019.105728>
- [2] A. M. Altabeeb, A. M. Mohsen, L. Abualigah, A. Ghallab, “Solving capacitated vehicle routing problem using cooperative firefly algorithm,” *Applied Software Computing*, vol. 108, e107403, Sep. 2021.
<https://doi.org/10.1016/j.asoc.2021.107403>
- [3] J. Cai, P. Wang, S. Sun, H. Dong, “A dynamic space reduction ant colony optimization for capacitated vehicle routing problem,” *Software computing*, vol. 26, no. 17, pp. 8745-8756, Sep. 2022.
<https://doi.org/10.1007/s00500-022-07198-2>
- [4] A. Thammano, P. Rungwachira, “Hybrid modified ant system with sweep algorithm and path relinking for the capacitated vehicle routing problem,” *Heliyon*, vol. 7, no. 9, e8029, Sep. 2021.
<https://doi.org/10.1016/j.heliyon.2021.e08029>
- [5] Y. Zhou, W. Li, X. Wang, Y. Qiu, W. Shen, “Adaptive gradient descent enabled ant colony optimization for routing problems,” *Swarm and Evolutionary Computation*, vol. 70, e101046, Apr. 2022.
<https://doi.org/10.1016/j.swevo.2022.101046>
- [6] P. Kalatzantonakis, A. Sifaleras, N. Samaras, “A reinforcement learning-Variable neighborhood search method for the capacitated Vehicle Routing Problem,” *Expert Systems with Applications*, vol. 213, e118812, Mar. 2023. <https://doi.org/10.1016/j.eswa.2022.118812>
- [7] A. Hottung, K. Tierney, “Neural large neighborhood search for routing problems,” *Artificial Intelligence*, vol. 313, e103786, Dec. 2022. <https://doi.org/10.1016/j.artint.2022.103786>
- [8] B. Kitchenham, P. Brereton, “A systematic review of systematic review process research in software engineering,” *Information and Software Technology*, vol. 55, no. 12, pp. 2049–2075, 2013.
<https://doi.org/10.1016/j.infsof.2013.07.010>

- [9] A. García-Holgado, S. Marcos-Pablos, F. J. García-Peñalvo, "Guidelines for performing systematic research projects reviews," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol.05, e05, 2020. <https://doi.org/10.9781/ijimai.2020.05.005>
- [10] W. Yang, L. Ke, "An improved fireworks algorithm for the capacitated vehicle routing problem," *Frontiers in Computer Science*, vol. 13, no. 3, pp. 552-564, Jun. 2019. <https://doi.org/10.1007/s11704-017-6418-9>
- [11] Y. Wu, F. Pan, S. Li, Z. Chen, M. Dong, "Peer-induced fairness capacitated vehicle routing scheduling using a hybrid optimization ACO-VNS algorithm," *Software computing*, vol. 24, no. 3, pp. 2201-2213, Feb. 2020. <https://doi.org/10.1007/s00500-019-04053-9>
- [12] Y. H. Jia, Y. Mei, M. Zhang, "A Bilevel Ant Colony Optimization Algorithm for Capacitated Electric Vehicle Routing Problem," *IEEE Transactions on Cybernetics*, vol. 52, no. 10, pp. 10855-10868, Oct. 2022. <https://doi.org/10.1109/TCYB.2021.3069942>
- [13] N. Frias, F. Johnson, C. Valle, "Hybrid Algorithms for Energy Minimizing Vehicle Routing Problem: Integrating Clusterization and Ant Colony Optimization," *IEEE Access*, vol. 11, pp. 125800-125821, 2023. <https://doi.org/10.1109/ACCESS.2023.3325787>
- [14] N. Sultana, J. Chan, T. Sarwar, A. K. Qin, "Learning to Optimise Routing Problems using Policy Optimisation," in *Proceedings of the International Joint Conference on Neural Networks*, 2021. <https://doi.org/10.1109/IJCNN52387.2021.9534010>
- [15] J. Li et al., "Deep Reinforcement Learning for Solving the Heterogeneous Capacitated Vehicle Routing Problem," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 13572-13585, Dec. 2022. <https://doi.org/10.1109/TCYB.2021.3111082>
- [16] Y. Wu, W. Song, Z. Cao, J. Zhang, A. Lim, "Learning Improvement Heuristics for Solving Routing Problems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 9, pp. 5057-5069, Sep. 2022. <https://doi.org/10.1109/TNNLS.2021.3068828>
- [17] Y. Xu, M. Fang, L. Chen, G. Xu, Y. Du, C. Zhang, "Reinforcement Learning with Multiple Relational Attention for Solving Vehicle Routing Problems," *IEEE Transactions on Cybernetics*, vol. 52, no. 10, pp. 11107-11120, Oct. 2022. <https://doi.org/10.1109/TCYB.2021.3089179>
- [18] K. Lei, P. Guo, Y. Wang, X. Wu, W. Zhao, "Solve routing problems with a residual edge-graph attention neural network," *Neurocomputing*, vol. 508, pp. 79-98, Oct. 2022. <https://doi.org/10.1016/j.neucom.2022.08.005>
- [19] Y. Hou, W. Gu, C. Wang, K. Yang, Y. Wang, "A Selection Hyper-heuristic based on Q-learning for School Bus Routing Problem," *International Journal of Applied Mathematics*, vol. 52, no. 4, pp. 817-825, Dec. 2022.
- [20] W. Yi, R. Qu, "Automated design of search algorithms based on reinforcement learning," *Information Sciences (NY)*, vol. 649, e119639, Nov. 2023. <https://doi.org/10.1016/j.ins.2023.119639>
- [21] H. Liu, Z. Zong, Y. Li, D. Jin, "NeuroCrossover: An intelligent genetic locus selection scheme for genetic algorithm using reinforcement learning," *Applied Software Computing*, vol. 146, e110680, Oct. 2023. <https://doi.org/10.1016/j.asoc.2023.110680>
- [22] A. Bogrybayeva, T. Yoon, H. Ko, S. Lim, H. Yun, C. Kwon, "A deep reinforcement learning approach for solving the Traveling Salesman Problem with Drone," *Transportation Research Part C: Emerging Technologies*, vol. 148, e103981, Mar. 2023. <https://doi.org/10.1016/j.trc.2022.103981>

- [23] W. Yi, R. Qu, L. Jiao, "Automated algorithm design using proximal policy optimisation with identified features," *Expert Systems with Applications*, vol. 216, e119461, Apr. 2023.
<https://doi.org/10.1016/j.eswa.2022.119461>
- [24] W. Yi, R. Qu, L. Jiao, B. Niu, "Automated Design of Metaheuristics Using Reinforcement Learning Within a Novel General Search Framework," *IEEE Transactions on Evolutionary Computation*, vol. 27, no. 4, pp. 1072-1084, Aug. 2023. <https://doi.org/10.1109/TEVC.2022.3197298>
- [25] D. H. Lee, J. Ahn, "Multi-start team orienteering problem for UAS mission re-planning with data-efficient deep reinforcement learning," *Applied Intelligence*, vol. 54, no. 6, pp. 4467-4489, Mar. 2024.
<https://doi.org/10.1007/s10489-024-05367-4>
- [26] Y.-E. Hou, W. Gu, C. Wang, K. Yang, Y. Wang, "A Selection Hyper-heuristic based on Q-learning for School Bus Routing Problem," *International Journal of Applied Mathematics*, no. 4, e35, Dec. 2022.
- [27] N. Furian, M. O'Sullivan, C. Walker, E. Çela, "A machine learning-based branch and price algorithm for a sampled vehicle routing problem," *OR Spectrum*, vol. 43, no. 3, pp. 693-732, Sep. 2021.
<https://doi.org/10.1007/s00291-020-00615-8>
- [28] A. Yaddaden, S. Harispe, M. Vasquez, "Transfer Learning Helpful for Neural Combinatorial Optimization Applied to Vehicle Routing Problems?," *Computing and Informatics*, vol. 41, pp. 172-190, 2022. https://doi.org/10.31577/cai_2022_1_172
- [29] M. Ben Hassine, M. S. Amri Sakhri, M. Tlili, "Machine learning approach to solving the capacitated vehicle routing problem," in *1st IEEE Afro-Mediterranean Conference on Artificial Intelligence, AMCAI 2023 - Proceedings*, 2023. <https://doi.org/10.1109/AMCAI59331.2023.10431502>
- [30] J. Dornemann, "Solving the capacitated vehicle routing problem with time windows via graph convolutional network assisted tree search and quantum-inspired computing," *Frontiers in Applied Mathematics and Statistics*, vol. 9, e1155356, 2023. <https://doi.org/10.3389/fams.2023.1155356>
- [31] D. Weiss, K. Tierney, "Realtime gray-box algorithm configuration using cost-sensitive classification," *Annals of Mathematics and Artificial Intelligence*, vol. 23, e9890, 2023.
<https://doi.org/10.1007/s10472-023-09890-x>
- [32] C. Wang, Z. Yu, S. McAleer, T. Yu, Y. Yang, "ASP: Learn a Universal Neural Solver!," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 6, pp. 4102-4114, Jun. 2024.
<https://doi.org/10.1109/TPAMI.2024.3352096>
- [33] M. Morabit, G. Desaulniers, A. Lodi, "Learning to repeatedly solve routing problems," *Networks*, vol. 83, no. 3, e22200, 2024. <https://doi.org/10.1002/net.22200>