Journal of Robotics and Control (JRC)
Volume 6, Issue 1, 2025
ISSN: 2715-5072, DOI: 10.18196/jrc.v6i1.24097

Machine Learning Paradigms for UAV Path Planning: Review and Challenges

Anis Mahmoud Bacha [1*], Razika Boushaki Zamoum [2], Fadhila Lachekhab [3]
[1,2,3] Laboratoire d'Automatique Appliquée, Faculté des Hydrocarbures et de la Chimie, Université M'hamed Bougara de Boumerdes, Avenue de l'Indépendance, 35 000, Boumerdès, Algérie
Email: [1] b.mahmoud@univ-boumerdes.dz, [2] r.boushaki@univ-boumerdes.dz,
[3] f.lachekhab@univ-boumerdes.dz
*Corresponding Author

Abstract—**Path planning is a crucial step in robotic navigation to satisfy: tasks safety, efficiency requirements and adapt to the complexity of environments. Path planning problem is particularly critical for Unmanned Aerial Vehicles (UAV), being increasingly involved within important tasks in diverse military and civil fields such as: inspection, search and rescue and communication, taking advantage of their high flexibility, maneuverability and cost-effective solutions. This continuous growth made the solution of UAV path planning problem an interesting research topic in recent years. In this scope, machine learning algorithms were a promising tool due to their continuous data-driven self-improvement to adapt with the high dynamicity of environments where conventional programming fails. This paper provides a review on recent developments in machine learning-based UAV path planning issued from credible databases like: IEEE, Elsevier, Springer Links and MDPI. The main contribution of this paper is to delve through these recent works providing a taxonomy of algorithms into the fundamental paradigms: supervised, unsupervised and reinforcement, evaluating their efficiency and limitations under distinct scenarios. Despite the relative generalization of deep reinforcement learning to different environments, this study highlighted some active challenges about computational cost and real-time applications that remain open.**

***Keywords*—UAV; Path Planning; Path Optimization; Machine Learning; Autonomous Navigation; Supervised Learning; Reinforcement Learning; Unsupervised Learning; Deep Learning**

I. INTRODUCTION

Unmanned Aerial Vehicle (UAV) or the so called *drone* offers a cost effective data exchange solution [1] resulting a fast grow in its use in various civil applications [2] after being restricted to the military domain where it made its first appearance[3]. Many of nowadays applications rely on drones: agriculture, meteorology, coastal surveillance, disaster management... with a promising growth in markets like: delivery, internet and telecommunication...[4]. In agriculture, UAVs enhanced precision and process monitoring of many activities like irrigation and optimized operational costs [5]. UAVs enhanced also safety of inspection missions in infrastructures especially thanks to their high accessibility to hard locations, limiting the risks on human inspectors and replacing traditional expensive equipments [6]. In disasters, drones offer a rapid assessment which accelerates emergency plans and search and rescue missions by the use of heat cameras [7]. These are only a sample example of tasks, many other applications solicited UAVs of their proven efficiency especially in critical tasks [2].

In parallel with this significant growth in UAV applications, the challenge to ensure a safe and efficient navigation in its environment rises, and hence the so called path planning problem. Also known as mission planning in military applications [8], this problem consists of guiding a UAV from a start to an end point in a given environment through a collision free path [9]. Obstacles can be basically static with a prior knowledge on the environment, however in real world path planning faces the dynamics of both unknown environment and the navigating robotic system as well [9]. Since the UAV is solicited for critical tasks in difficult terrain, path planning problem deals with several challenges. In delivery tasks for example, UAV is said to traverse a highly dynamic and safety demanding environment, so real-time re-planning ability is a must, with the consideration of completion time that impacts directly the productivity [10]. UAV is also physically limited by energy consumption and CPU-cost. Resource management is then an important constraint for path planning [11].

The importance of path planning problem triggered an evolution of solution algorithms from simple deterministic to complex adaptive [12]. The growing complexity of environment amplified the need to autonomous real-time solutions and learning based approaches[12]. Traditional methods commonly fail to complete the tasks in the dynamically-changing scenarios [13] . Machine learning algorithms represents a promising tool for the purpose of finding an optimal path for a UAV. In contrast to classical programming which relies on programmer abilities to predict all cases [14], machine learning uses data subsets to train computer which learns itself and adapts itself to newer datasets [15]. The fusion of machine learning and UAV has brought many advantages: it enhanced decision making process following a data-driven autonomous approach, improved adaptability to dynamic environments and ensured a continuous performance improvement through learning [16].

UAV path planning solutions had been the subject of interest for many reviewers who differ in their specific focus. Most authors like: Ajith and Jolly [17], Hooshyar and Huang [18] and Yahia and Mohammed [19] focus on metaheuristic techniques as being more popular in UAV subjects. The authors analyzed many optimization techniques to path planning constraints but excluded machine learning alternatives. Coming to Ait Saadi et al. [20] who classified previous planning algorithms into: classical, heuristic, metaheuristic, machine learning and hybrid. Advantages and disadvantages of each class were summarized regarding the results of reviewed articles. The authors succeed to draw some common features and weaknesses of each class of algorithm where machine learning algorithms were proven to be highly adaptive to abrupt changes but costly and data-hungry. The article lacks to delve into subcategories of each class and does not explore sufficiently the differences within machine learning paradigms. Similarly, Puente-Castro et al. [8] reviewed UAV swarm planning approaches using neural networks, reinforcement learning and metaheuristic algorithms. The authors widely covered various classes of algorithms but did not include all machine learning paradigms, especially unsupervised learning which is useful in multiple UAV systems enhancement. The rising interest in learning-based algorithms was represented by Zhang et al. [21] who explored deep reinforcement learning methods and delved into performance improvements to be applied to the network and exploration-exploitation balance. On the other hand, supervised and unsupervised methods were missing in this proposal.

In comparison with the previously stated existing review papers, the main contributions of this proposal are:

- Review recent UAV path planning works that resorted to machine learning algorithms within paradigm taxonomy: supervised, unsupervised and reinforcement.
- Evaluate their performance and deduce the suitable scenarios and constraints of each paradigm in path planning.
- Discuss advantages and drawbacks of each technique.
- Analyze improvement tools and their efficiency.
- State the limitations and highlight future directions.

The organization of this paper is outlined as follows: Section II explains the methodology of the paper and proposal's selection criteria, brief definitions of common concepts are stated in SectionIII then the literature review is conducted along Section IV. Section V comments the findings of its previous section. Finally, Section VI closes the review with conclusions and future orientations.

II. Paper Methodology

This paper consists in a review of existing UAV path planning approaches that relied on machine learning algorithm. This work is conducted as shown in Fig. 1. After defining the motivations behind this work and its objectives, an overview of common concepts is stated prior to the survey, in order to familiarize with its topic. Based on common results gathered from selected papers, an analysis is conducted to discuss main findings and highlighting bibliometric indexes. Finally, the paper is concluded by summarizing findings and limitations, and proposing future insights.

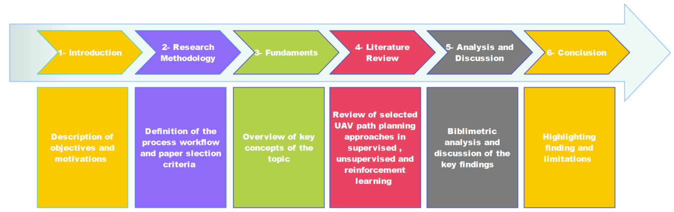

Fig. 1. Process workflow of the paper

The search for recent proposals in machine learning based UAV path planning was done through *Google Scholar* engine. Several search trials were done using different string entries but all convey to the following research equation:

”*ML_approach_name*” **AND** ”UAV” **AND** (”*path planning*” **OR** ”*navigation*”)

The selection process of papers is demonstrated through the flowchart shown in Fig. 2. The selected papers are said to obey some studied criteria. The proposal should satisfy the research equation and it should be dated posterior to 2018. Nevertheless, few contributions prior to 2018, or with mobile robot as agent are accepted in case the used algorithm is not already sufficiently covered by papers satisfying the criteria. The selected proposals must respond clearly to key research questions: which algorithm ? which agent ? which environment (scenario) ? what are the comparative results ?. These questions are helpful to build an appropriate evaluation.

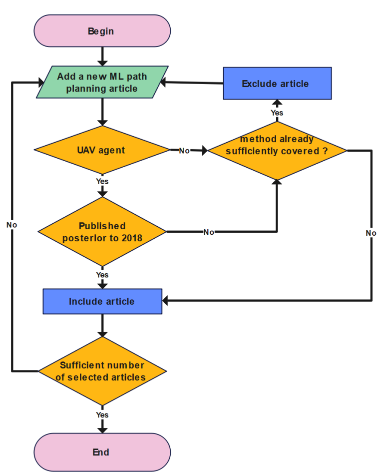

Fig. 2. Papers selection process

III. Fundaments

A. Unmanned Aerial Vehicles (UAV)

Generally called *drones*, unmanned aerial vehicles (UAV) are pilotless aircrafts [22] that flies autonomously or can be controlled remotely, all relying on computer and electronic subsystems [8]. Drones, which were limited to military tasks, are currently receiving an increasing demand from various commercial and industrial fields registering important market value in[23][24]: agriculture, delivery, infrastructure, media and others.

UAVs can be classified regarding different parameters. The wing configuration design is a fundamental attribute which decides the way of taking-off and landing. Two types are distinguished, Vertical Take-Off & Landing (VTOL) and horizontal take off landing (HTOL)[25]. Wing configuration has an important impact on path planning, rotary-wing drones are highly maneuverable and suitable for quick changes, whereas fixed-wing UAVs are more sensitive to sharp turns. Hassanalian and Abdelkefi. [26] proposed also a classification based on size and weight of a UAV. The authors presented a spectrum spread of UAVs from weight a 15,000 Kg and a wing span of 61 m to the tiny category called Smart Dust (SD) whose weight is less than 0.5 g and its wing span does not exceed 2.5 mm. Watts et al. [27] proceeded into a classification with respect to their attitude and flight endurance where the authors distinguished: Low-Altitude-Short-Endurance (LASE) where an attitude of 1,500m is not exceeded, LALE Low-Altitude-Long-Endurance (LALE) with extended flight period, Medium-Altitude-Long-Endurance (MALE) which they can fly at an attitude of 9,000m for many hours and High-Altitude- Long-Endurance (HALE) with an attitude range of 20,000m and an endurance up to 30 hours. These classifications are helpful in the choice of the appropriate UAV regarding task constraints: LALE UAVs have sufficient capacities for delivery tasks, where military missions may require HALE drones.

B. Path Planning

Path planning problem can be defined briefly and thoroughly as the process of seeking for an optimal obstacle-free path for a mobile robotic system from a starting point (initial state) to a target point (goal state) in a given environment, using accumulated sensor data and initial information [28, 29]. Puente-Castro et al.[8] devised the process of path planning into two elementary steps: environment modeling and path search.

Environment modeling is a fundamental step in many path planning approaches and there are different ways to model an environment [30]. Han et al.[31] gave a classification of environment modeling methods into geometric, graph and grid based methods. The geometric approach represents obstacles by geometric figures and their interconnections [32]. Graph theory is the basis of graph methods where airspace is abstracted into a graph, Zhang et al.[33] used a Voronoï graph representation. The grid based representation divides in an efficient manner a 3D or 2D space into uniformly sized cells [34, 35]. Nowadays many researches bypassed the environment modeling step by using popular high-fidelity simulation platforms like ROS-Gazebo [36] and Microsoft Airsim [37] which are open source and under continuous development to mimic real-world dynamics. These tools handles training and the development of path planning learning models through real-time simulation in a cost-effective manner.

Path search and optimization techniques were investigated and classified by Reda et al. in [38]. The first category included classical approaches such as graph based, sampling based, artificial potential field and curve fitting techniques. These traditional methods had shown some limitations. Graph based methods give jerky paths in large-scale environments, sampling based approaches are faster but fail into the same limitation. Gradient based methods are time-efficient collision free planners but commonly fail in local minimum. Curve fitting techniques are high computation cost methods. The second category defined by Reda et al. [38] was metaheuristic class of algorithms that are the most popular path optimization techniques due to its applicability to several scenarios. Nevertheless, this class does not guarantee the convergence to optimal solutions. The third category was machine learning algorithms, the topic of interest in this paper. Learning-based algorithms offer a high level of autonomy and enhance optimization by reducing navigation sensors number and hence battery cost [8]. Machine learning algorithms are high speed solutions in familiar scenarios, main common challenge of this class is to faster training time in unknown scenarios [38].

Regarding available information about the environment, path planning problem is divided into two classes [28, 39]. Global path planning also known as offline [30] where the environment is static and all information about it known before the first move. Local path planning, called also on-line [40] has a dynamic environment and its information is partially or completely unknown [41].

Ait Saadi et al. [20] summarized some objectives and constraints of path planning task. The objectives can be: path smoothness, time-efficiency, cost (CPU, battery) optimization and collision avoidance. The constraints on the other hand can be seen as limitations in attitude, velocity, battery and computational performance, also environment obstacles and threats. These constraints significantly impact algorithm performance, the simplest illustration is the consideration of obstacle-avoidance in a dense complex environment will induce a relative increase in computation cost [42] [43].

C. Machine Learning

Machine learning (ML) is considered as a field of computer science having the aim to automate the solutions to problems which are hard to solve using classical programming [44]. In other words, a process that enables computers to learn and improve on their own [45]. Machine learning algorithms are classically divided into three main paradigms (Fig. 3): supervised learning, unsupervised learning and reinforcement learning [46].

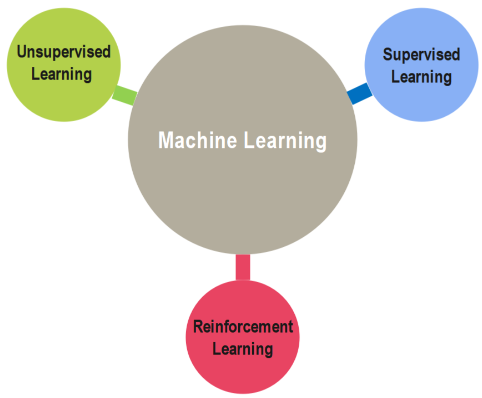

Fig. 3. Machine learning paradigms

Supervised learning consists of labeled data training in order to end up with a mapping (model) between input (data) and output (labels), this model will be used in the prediction of output of unknown data set [47]. Supervised learning method is itself divided into two sub-categories: regression where the output is continuous, and classification where the output is discrete [48].

Contrary to the first paradigm, unsupervised learning deals with unlabeled input data where there is no explicit output associated. The aim of unsupervised learning is to group and represent a given input pattern in a convenient way that matches the overall structure of patterns [49]. Clustering algorithms are the famous class of this paradig where items are grouped through clusters based on defined similarities [50].

Reinforcement, the third paradigm, also does not require labeled data, but uses the interaction of the agent in an environment, and learns from this experience. This can be represented by Markov Decision Process (MDP): an environment reflected by an actual state, the agent performs a given action changing that state and getting a resulting reward [51].

IV. Literature Review of Recent ML Based Path Planning Algorithms

A. Supervised Learning

Regression analysis consists of fitting a right model (mapping) matching one dependent variable (target) to one or more independent variables (predictors), and based on the obtained model the machine makes a prediction for future output [52]. Meng et al. [53] relied on a linear regression (LR) model to design a GPS anti-spoofing system for a UAV. The linear regression model was used to predict the 2D trajectory of UAV whose coordinates are then compared to the GPS actual positioning. A threshold difference was fixed to decide whether the UAV was cheated and corrective destinations were required. The linear regression in this point to point traveling problem has shown by far a higher accuracy than Long Short Term Memory Recurrent Neural Network (LSTM-RNN).

Kumar et al.[54] used a multiple linear regression (MLR) to plan a Nao humanoid robot facing static and dynamic obstacles. First in a static environment, regression algorithm was trained by distances to three obstacles: in front, at the right side and at the left side of the robot, to give the turning angle as an output. The author added two Naos humanoid robots as dynamic obstacles and enhanced the regression analysis with a six (6) positions Petri-net priority controller within each robot where the shorter the distance to the target, the higher the priority. The proposed method was tested in both simulation and experimental platforms considering two cases: a unique Nao robot with static obstacles , and three Naos simultaneously. Simulation results has shown about a 6% shorter path than Co-Evolutionary improved Genetic Algorithm (CEGA) and an improved Genetic Algorithm (IGA) developed by Qu et al. [55].

A polynomial regression (PR) path planning for UAV was proposed by Koo et al. [56] in a wireless sensor network (WSN). The challenge was to allow the UAV to cover all sensor nodes; hence, a modified Mean Squared Error (MSE) function taking into account transmission range was proposed in a double loop path planning process. The inner loop generates the path considering the distance to sensor's nodes, the outer loop refines the path considering the transmission range of the sensors. The simulation resulted a 10% shorter path than Genetic Algorithm (GA) and enabled the UAV ensure all data transfer.

Boulares and Barnawi [57] proposed a novel floating target search algorithm for a UAV in ocean environment using Lasso regularized polynomial regression (L-PR). The authors applied first a recursive clustering process to define strategic zones based on drifting history. The weight centers of the clusters are smart searching areas for the UAV. Then, a particle trajectory simulation was applied taking into account wind and ocean currents forces parameters. Finally enhanced polynomial regression approximation is applied to acquire a set of predicted paths around each weight center. Through simulation, the authors proposed three regularization techniques for polynomial regression model to overcome ordinary least squares (OLS) estimation limitations. The three were: Lasso regularization (L-PR), Ridge regularization (R-PR) and Elastic net regularization (EN-PR). The results were evaluated regarding R^2 metric. Ridge regularization with a polynomial degree of 26 registered the best score among the three approaches.

A Gaussian process regression (GPR) was suggested by Yel and Bezzo in [58] to plan a safe path for a quadrotor UAV under payload change disturbance. The solution consisted of an offline stage where the GP was trained to estimate the payload and the trajectory deviation, and an online stage to adapt the GP model with real time data and determine the adequate speed. The approach was validated in simulation and experiments in static environments with window obstacles.

Support vector machine (SVM) is a powerful solution in binary classification problems. SVM maps training data vector to a space of higher dimension, separating the two clusters by a linear or non-linear kernel function [59]. Al-Naeem et al. [60] proposed an energy efficient path planning algorithm for a UAV in a large agricultural field based on SVM. The UAV system was deployed to collect data from Iot (internet of things) devices. SVM was used to classify agricultural sub-regions into sensitive and insensitive ones based on a set of collected and stored attributes. The sensitive sub-regions, where most anomalies were detected recently, had then more frequent UAV visits. The simulation resulted a better performance of this solution than A* Genetic Algorithm regarding successful detection rates, time delay and energy consumption.

Chen et al. [61] applied a Gaussian kernel SVM to plan a smooth path for a UAV in surface of minimum risk (SOMR). An offline 3D SOMR data were obtained apriori defining flyable and no-fly zones, start and goal points. The SOMR was squeezed into a simplified 2D safe-map. Then, online SVM was applied to divide obstacle points into two (2) classes. Virtual obstacle were distributed around starting and arrival points to maintain them within the optimal path. Finally, the 3D optimal path in the SOMR was implemented by correspondence. Simulation results has shown the feasibility of the approach.

On the other hand, a multi-class SVM (MSVM) algorithm was proposed by Morales et al. [62] as a solution for mobile robot path planning problem in a dynamic environment. In this proposition, SVMs took points of an aerial image map where all obstacles were represented as input, and the one-versus all using a winner-takes-all multi-class strategy was adopted to extract a safe path for each obstacle. The significant change in dynamic patterns only affects the pattern itself and not the whole map. The chosen optimal path at the end of the process joins feasible paths intersections. This solution was tested in both simulation and experiment using 4 different maps (environments). The MSVM algorithm has shown a higher performance than single class SVM, Voronoï diagram and A* algorithms in terms of path length, smoothness and distance to obstacles, with a full success rate.

Asti et al. [63] used K-Nearest-Neighborhood (KNN) classifier in the design of a solution to UAV path planning problem with static obstacles avoidance. The key contribution of KNN was to select the direction of obstacle avoidance between three choices: up, right and left. The selection was based on the magnitude of the deviation vector and energy consumption during deviation process. The approach was trained and tested in three static environment with 1, 2 and 3 obstacles respectively resulting 96.6% accuracy, 0.0068 s time cost and 0.64 m/s minimum velocity.

Pandey et al. [64] had a wheeled robot path planning problem in a static environment, to which he proposed an FNN based solution, tuned by particle swarm optimization (PSO) algorithm in order to minimize the error. FNN receives distance to obstacles as input and outputs the steering angle. Simulation results qualified this approach to be more cost effective regarding computation time and path length compared to an FNN without tuning and fuzzy-PSO methods.

Sanna et al. [65] proposed an FNN solution to coverage path planning (CPP) by multiple UAVs. The authors proceeded first to divide the coverage grid area using k-means clustering into sub-areas where a single UAV is assigned at each sub-area centroid. UAV had an 9X9 cells field of view inside a 19x19 local field and it is said to visit a non-obstacle cell only once. For this purpose, ANN was trained with 54˙531 labeled data with an accuracy of 92%. After the coverage of the local view Explorative A* and A* to visit the nearest unexplored cell considering other agents avoidance. Simulation of the proposed A*-ANN was conducted through four scenarios one geometric map and three real occupancy grids and the number of UAV varied from 3 to 6. The results has shown that increasing the number UAVs reduced the number of moves.

Choi et al. [66] has proposed a solution for an indoor UAV path planning with climbing stairs option relying on a CNN. The CNN was used precisely as an image processing system to identify stairs and a Light Detection and Ranging (Lidar) sensor was the solution for distance measurement. The results was a collision-free stair climbing with 92.06% accuracy in stair recognition.

In the proposal of Liu et al. [67], the authors designed a UAV residual path planner (Res-planner) based on residual CNN (R-CNN). The methodology started by sampling offline scenarios to get state-behavior couple at each time t. These couples were fed as training datasets into the R-CNN network and then predicted behavior was constructed. In the testing phase the R-CNN received sampled state containing information about distance and obstacles and outputs the heading direction (behavior). In 200 static scenarios, R-CNN obtained 88.2% feasible path and 72% optimal. In local planning scenarios. R-CNN outperformed A* in terms of time cost and path length. The authors proved also through experiments that increasing the training datasets number (the number of samples) and the number of the network layers increases the accuracy and performance of R-CNN approach.

Dai et al. [68] proposed a CNN-based path planning for a quadrotor UAV to avoid obstacles in an unknown environment. The proposed CNN model received a monocular forward-facing camera image. Based on a confusion matrix, the network predicted the collision probability and the needed steering angle. The two prediction were inputs to the control mechanism acting on the yaw angle and forward speed. The response speed in state transitions were enhanced by the mean of a first order butterworth filter. The model was trained using large existing and experiment datasets of frames. The approach was tested through 5 different indoor/outdoor real scenarios varying in complexity and lighting conditions. The model offered a low sensor requirement obstacle avoidance technique, performing better with gray-scale images than RGB images regarding flight distance.

Sartori et al. [69] used CNN for the estimation and refinement of an indoor global path planning for a mobile robot. The CNN received a top image of the 2D map defining the start and target points and obstacle distributions to provide the path way-points' coordinates as output. This described approach resulted a cost effective safe path planning solution in three realistic indoor environments with a significantly high success rate.

Nair and Supriya [70] proposed an RNN solution for a mobile robot path planning in a dynamic environment. Precisely, a Long Short Term Memory (RNN-LSTM) network was designed and trained in 12 different environment. Despite of some failed outputs, simulation results has shown a higher efficiency in obstacle avoidance of this approach compared to A* algorithm.

The main results issued from the reviewed supervised learning algorithms are summarized in Table I.

B. Unsupervised Learning

Based on similarities, K-means clustering algorithm divides a dataset into K groups (clusters) each cluster have a centroid which will be updated when we introduce a new item; the new item is assigned to the cluster having the closest centroid [50]. Yue and Zhang [71] proposed a k-means clustering algorithm supporting a simulated annealing (SA) algorithm to plan the path of multiple UAVs in a cruise coverage mission. After defining the valid flyable cruise area that meets the UAVs constraints, K-means algorithm was established to cluster target points then the total cruise area was divided according to the number of UAVs. The simulation resulted a successful coverage of more than 92% of the area with 30 UAVs, and the authors concluded that this combination of SA-K-means algorithm performs better than genetic algorithm (GA) and Hop-field Neural Network (HF-NN) regarding convergence to global optimal solution.

In [72], Li et al. also used k-means clustering to support an improved ant colony algorithm as a solution to UAV path planning for navigation mark inspection mission. The idea was to divide inspection area to reduce traveling time; thus, K-means algorithm was used to group mark inspections into distinct clusters where UAV will inspect a single cluster per flight. Both experiment and simulation has proven the efficiency of K-means algorithm which reduces flight range by 20% compared to ant colony algorithm alone, after dividing the inspection marks into three (3) clusters.

An improved K-means (I-K-means) algorithm with three-stage clustering was proposed in [73] as a support to Way-point Refinement Iteration (WRI) path planning algorithm for multiple UAVs in a WSN. The idea was to cluster first the sensors, then select a cluster head (CH) sensors. These CHs sensors collects all data from sensors belonging their clusters. Finally those CH sensors are themselves clustered by the number of UAVs and thus, the mission is assigned. The simulation considered 500 sensors and 4 UAVs in two scenarios: a balanced and an unbalanced sensor distribution. The results shown the relatively higher efficiency of k-means clustering compared to Salp-Swam optimization (SSO) proposed in [74] with 26 %, and 7% shorter average distance from the cluster center in scenario 1 and scenario 2 respectively. Furthermore a 23% more efficient path was registered by the combination of WRI and K-means (WRI-I-K-means).

Similarly, Li et al. [75] resorted to K-means clustering in their proposal for multiple UAVs path planning in a WSN. The authors proceeded into clustering sensor nodes into k clusters, where k is the number of deployed UAVs. An improved MIN-MAX Ant System (improved MMAS) algorithm was used to optimize the path and solve local optima problem. The approach was simulated in a scenario with 40 sensor nodes and 3 UAVs. The proposed K-means-Imroved MMAS algorithm resulted a shorter path compared to original MMAS and the local optima is avoided after 80 iterations.

Ma et al. [76] introduced k-means clustering as an enhancement to GA in the solution of multi-UAVs task assignment and path planning. The problem was defined as M-TSP problem where multiple UAVs are said to visit each way points and return to starting points in a closed paths. The proposed coordinated optimization algorithm combining GA and K-means takes into consideration the maximum flying distance in the selection of the clusters number. In simulation the number of tasks to be assigned was 100 and the flyable distance was set to 50Km, 40Km, 30 Km and 20Km. The number of UAVs (clusters) was found to vary inversely to maximum flight distance. Despite additional time cost, the proposed approach outperformed earlier GA with respect to path length in all scenarios.

Suseno and Wardana [77] suggested a two stage progressed K-means (P-K-means) clustering to a UAV path planning problem in a maritime surveillance mission. The progressed clustering method does not define a fixed number of clusters as input, but rather it defines a maximum radius. The first clustering stage was to divide the mission into operation areas based on the ships distribution. The second stage clustering was a smaller scale clustering aiming to select mini-clusters at each operation area to group adjacent ships together through

TABLE I. SUPERVISED LEARNING PROPOSALS FOR UAV PATH PLANNING

Algorithm	Author(s)	Agent	Environment	Main Result	Year
LR	Meng et al.[53]	UAV	2D, dynamic	higher prediction accuracy than RNN-LSTM	2021
MLR	Kumar et al.[54]	mobiile robot	2D, static & dynamic	shorter path compared to CEGA and IGA	2018
PR	Koo et al. [56]	UAV	2D static	shorter path compared to GA	2020
L-PR	Boulares and Barnawi [57]	UAV	2D dynamic	the best 26 degree polynomial R^2 accuracy score compared to R-PR and EN-PR	2021
GPR	Yel and Bezzo [58]	UAV	3D static	successful obstacle avoidance	2020
SVM	Al-Naeem et al. [60]	UAV	2D	optimal time and energy efficient completion	2023
G-SVM	Chen et al. [61]	UAV	3D static	smooth obstacle-free path	2014
MSVM	Morales et al. [62]	UAV	2D, static & dynamic	better performance than Voronoï and A* in terms of path length and smoothness and distance to obstacles	2016
KNN	Asti et al. [63]	UAV	3D, static	96% accurate obstacle avoidance	2020
FNN-PSO	Pandey et al. [64]	mobile robot	2D, static	better performance in terms of path length and computation cost with respect FNN alone and fuzzy-PSO	2020
FNN-A*	Sanna et al. [65]	multiple UAVs	2D, static	92% training accuracy and better path efficiency with increasing UAVs number	2021
CNN	Choi et al. [66]	UAV	3D, static	92% accuracy and collision-free mission	2021
R-CNN	Liu et al. [67]	UAV	2D static & dynamic	equal optimal path with A* in static scenarios and a better computational cost in dynamic scenarios	2022
CNN	Dai et al. [68]	UAV	3D, static	low sensor requirement and better performance in terms of distance with gray scale images compared to RGB	2023
CNN-Theta*	Sartori et al. [69]	mobile robot	2D static	high success rate and low computation cost	2021
RNN-LSTM	Nair and Supriya [70]	Mobile robot	2D, dynamic	shorter path compared to A*	2020

the definition of weighted vulnerable points (VP) for each mini-cluster. Nearest neighborhood algorithm was designed to seek an optimal path joining the VPs. Simulation results in 50 Km, 100 Km and 150 Km operation area radii has shown that the number of operation areas decreases with the increase of their radii. In some scenarios the path was not optimal, but in terms of computation cost, this proposal was efficient compared to ant colony (ACO) and Held & Karp algorithms.

Chen et al. [78] proposed a density based spatial-temporal clustering algorithm (STCA) of regions in a coverage path planning problem using heterogeneous UAVs. The clusters centers were identified regarding higher density and farther distance to other high densities. The task assignment is done regarding the flight time and speed. The battery level is also taken into account in the decision of adding queued region to a current task charged UAV through two different strategies: nearest to end (head or tail regions) of an allocated cluster (STCA-NE), and nearest to any region in an allocated a cluster. Genetic algorithm GA is added to optimize the order of regions. Simulation results has shown that STCA-NE-GA (nearest to end with genetic algorithm) method had a better performance regarding completion time with 3% less than STCA-NA-GA (nearest any with genetic algorithm).

Similarly, multi-region coverage path planning problem of heterogeneous UAVs was tackled by Xiao et al. in [79] by the mean of clustering. In this proposal, UAVs were said to cover rectangular regions. The authors proposed a coverage-based scanning clustering algorithm (CSCA) that takes as initial cluster centers the UAVs bases. The algorithm keeps updating the centers according to spatio-temporal similarities between clusters centers and region centers, and remaining flight endurance of UAVs. Nearest to end (NE) policy was adopted as regional sorting strategy, and bilateral shortest-selection strategy (BSSS) was used to seek the shortest scanning path of each region. The approach was simulated in a squared 50Km mission area, 8 UAVs and 5 to 40 regions with an increment of 5. CSCA outperformed STCA with up 21% less task completion time. NE strategy sorting resulted much lower time cost than genetic algorithm GA. BSSS scanning path planning resulted a shorter path than long edge scanning strategy (LESS).

Dai et al. [80] suggested a game based cluster head selection algorithm (CHSA) in order to reduce energy loss for multiple UAVs in delivery task assignment. In this proposal, after an area division process, representative nodes were selected and then cluster was chosen based on mixed game model that prioritizing nodes with the highest energy-distance to representative node ratio. 100 nodes were deployed in 100x100 m simulation environment. Results has shown that CHSA outperformed k-means and ACO in terms of nodes survival rate and energy consumption.

Faigl et al. [81] proposed a surveillance planning of multiple UAVs with the help of a self organizing map (SOM). The problem was defined as a Dubin traveling salesman problem with neighborhoods (DTSPN) where the vehicle is requested to visit a set of way points and return to to its initial location in a closed contour. The SOM is a single layer of weighted neural network. Its process started with neuron process starts by competitive search of the winner neuron closest to the target then its neighborhood neurons weights were adapted by getting closer to the winner neuron based on a neighborhood function. In a simple environment (22 targets), the performance of SOM based DTSPN was compared to variable neighborhood search (VNS) algorithm in scenarios of single and multiple (up to 3) UAVs. SOM had the best performance regarding

CPU cost. In more complex environments (up to 100 targets), the results has shown that SOM based Bézier outperformed earlier methods with respect to traveling time while SOM based DTSPN remains with the best CPU cost.

Principle component analysis (PCA) is an unsupervised ML tool aiming to decrease the size of a dataset conserving sufficient information. Kishimoto et al. [82] applied PCA to a mobile robot equipped with a gamma-ray detector in order to minimize the navigation distance to radiation sources. First, Compton camera measured the radiation. Then, the radiation distribution was reconstructed using back-projection onto 2D space. PCA was used to generate an inspection path surrounding the source of radiation whose amount was compared to a threshold to choose whether the next step follows the first principal component vector or the second. Experiments were conducted using a single and multiple radiation sources and localization error were 0.35m and 0.28m respectively in simulation. This method has shown a higher performance than information-driven solution in [83] with less localization errors and measurement points. The main results issued from the reviewed unsupervised learning algorithms are summarized in Table II.

C. Reinforcement Learning

Q-learning is one of the most applied RL algorithms and is behind the foundation of many other algorithms [85]. This algorithm is used to solve problems in environments with limited states and discrete actions [51]. Gao et al. [86] introduced a risk free Q-learning (RFQL) algorithm to prevent UAV from blind navigation and ensure safe path in large scale space. The authors proposed two improvements. First, the distance to the target was taken into consideration to guide the training. Furthermore, states close to obstacles were identified through training experience as risk zones. Simulation in a 2D environment has shown that RFQL generated a safer path than standard Q-learning and converges 600 steps faster reducing iteration cost.

In [87] Q-learning was employed by Xie et al. to optimize the distance and energy of a navigating fixed wing UAV. The environment consisted of a battlefield area where the UAV loads data its sources and offloads it at the command post, avoiding fire threats. The approach considered the optimization of: flight distance, flight energy, communication energy while hovering and processing energy. Simulation was conducted in a $2000m \times 2000m$ decomposed into 10×10 grids.This Q-learning based approach resulted a smoother and 26% shorter path than ant colony algorithm (ACO).

An improved Q-learning algorithm was proposed in [88] by Yan and Xiang to plan an optimal path of a UAV in a reconnaissance mission. The authors used ϵ-greedy strategy alongside with Boltzmann strategy in order to avoid state redundancy. Furthermore, the distance to the target was taken into account in the Q function initialization through an exponential decaying function of the coordinates. The simulation considered a 2D airspace established in STAGE software were the UAV is trained 25000 episodes. The results has shown that the proposed improved Q-learning converged 2300 episodes faster than classical Q-learning and remained stable. Also, the path obtained through the improved Q-learning was four (4) steps shorter.

SARSA stands for State–Action–Reward–State–Action is also like Q-learning used to solve problems in environments with limited states and discrete actions, but it updates after performing the next action without taking a greedy action like in Q-learning algorithm [51]. Boming et al. [89] proposed an enhancement to SARSA algorithm for a UAV path planning problem. The approach was based on an ϵ-greedy strategy guided by the geometric distance to the target. In a static 2D environment, simulation results has shown a faster convergence of the proposed improved SARSA than ordinary SARSA and Q learning algorithms. In addition, this improved SARSA algorithm resulted an 18 steps shorter path than Q learning and a 2 steps shorter path than ordinary SARSA.

Huo et al. [90] proposed a Dyna-Q method to plan the path of multiple quadrotor UAV. The learning approach of this technique combine model-based and model free methods, where the agent interacts directly with its environment obtaining a real experience, and the environment model gives an estimated experience in order to rationalize exploration. The approach was tested in 2D 20X20 grid environment through static then dynamic scenarios. The results has shown that UAV reached the target fast after 150 training episodes in the static scenario, and after 1500 training episodes in the dynamic scenario.

Cui and Wang [91] introduced a 2 layer Q (2L-Q) learning algorithm to tackle a UAV global and local path planning problem. The idea was to assign global and local path planning processes to two separated Q learning layers. The global layer was to be activated when no obstacle motion is detected, following ϵ-greedy and Boltzman strategies. The local layer considered the distance and the angle between obstacle and target in the reward. The action with best Q value among the two layers at the current state was to be executed. The approach was trained in 2D static and dynamic environment achieving 89.4% success rate after 800 episodes. The test were conducted in 5 2D simulation scenarios varying from simple static to complex dynamic. The higher the complexity the longer the optimal path got.

It is hard and computationally expensive to deal with a large state space using basic Q-learning and SARSA. A DNN is employed as an approximation function for Q function and relies on an experience replay process to stabilize this approximation. The whole architecture is known as Deep Q Network DQN [92]. In [93], Anas et al. compared DQN algorithm to basic Q-learning and SARSA as a solution to a mobile robot

TABLE II. UNSUPERVISED LEARNING PROPOSALS FOR UAV PATH PLANNING

Algorithm	Author(s)	Agent	Environment	Main Result	Year
SA-K-means	Yue and Zhang [71]	multiple UAVs	2D, dynamic	better convergence to global optima than GA with 92% successful coverage	2018
ACO-K-means	Li et al. [72]	UAV	2D, dynamic	20% less flight range tan ACO alone	2023
WRI-I-K-means	Kim and Park [84]	multiple UAV	2D, dynamic	shorter distance to the clusters centers than SSO	2023
MMAS-K-means	Li et al. [75]	multiple UAV	2D, dynamic	better performance than MMAS alone regarding path length and local optima avoidance	2018
GA-K-means	Ma et al. [76]	multiple UAV	2D, dynamic	better performance than GA alone regarding path length at the expense of additional CPU cost	2019
P-K-means	Suseno and Wardana [76]	multiple UAVs	2D, dynamic	less CPU cost compared to ACO and Held & Karp algorithms	2021
STCA-NE-GA	Chen et al. [76]	multiple UAV	2D, dynamic	3% better completion time than STCA-NA-GA	2021
CSCA-NE-BSSS	Xiao et al. [79]	multiple UAVs	2D, dynamic	NE strategy generated CSCA and NE got better completion time performance than STCA and GA respectively, BSSS resulted a shorter path than LESS	2023
CHSA	Dai et al. [80]	multiple UAVs	2D, dynamic	CHSA had the best performance regarding energy cost compared to K-means and ACO	2022
SOM-DTSPN	Faigl et al. [81]	multiple UAVs	2D, dynamic	better performance with respect to travelling time than VNS and less CPU cost than SOM-Bezier	2019
PCA	Kishimoto et al. [82]	mobile robot	2D dynamic	higher performance and accuracy than data-driven solution [83]	2021

path planning in a static environment. DQN was the unique algorithm to approach 100% success rate in both simulation and real implementation.

Luo et al. [94] proceeded to Deep SARSA to solve a multiple UAV path planning in a dynamic environment. Similar to DQN, this approach consists of using a DNN to predict the action rather than Q table in simple SARSA, when it comes to large state space. The authors used a simplified 2D environment to train the UAVs where the success rate increases after the 800 first episodes following the ϵ-greedy strategy. The algorithm was tested in a virtual 3D environment deisgned in Ros-Gazebo platform, where a scenario of two UAVs, one static obstacles and two dynamic obstacles (2 other UAvs) was considered. The approach resulted a collision free-path.

Wang et al. [95] proposed a deep double Q network (DDQN) algorithm to plan the path of a UAV in a rescue mission in mountain environment. DDQN uses two Q networks one for action selection and the other for action evaluation to overcome overestimation of DQN. The authors generated a grid based 3D mountain environment using peaks function modeling. The simulation considered two scenarios: static target and dynamic target. Considering static target DDQN had a 200 episodes faster convergence than DQN, thus DDQN completed the task quicker. In the dynamic target case, DDQN was the only algorithm to converge within the 1000 training episodes.

Yan et al. [96] introduced Dueling DQN (D3QN) for a UAV path planning in dynamic environments considering potential enemy threats. D3QN architecture is constituted by splitting the Q networks into two streams: one to evaluate the state value and the other for advantage estimation. The algorithm was trained in a simulated 2D static then dynamic environments designed in STAGE platform, ϵ-greedy strategy was set with further heuristic guidance for action selection. D3QN has shown a better learning stability and performance in the static scenario compared to DQN and DDQN, and a higher success rate in the dynamic scenario. The test scenario contained one static threat and two dynamic ones. UAV reached the target safely through D3QN planner with a success rate of 67.33% although the scenario was not seen during training.

Chao et al. [97] proposed an Event-Based DQN (E-DQN) to UAV under unknown environment. The approach relied on event camera pictures whose features were extracted through spatiotemporal decoupling and reconstructed at the output of an auto-encoder. The reconstructed image were fed to train a DQN with an experience replay. The ϵ-greedy strategy were chosen to manage exploration and exploitation trade-off. This proposed E-DQN method was tested in Airsim simulation environment with unknown obstacles within 40 000 episodes. The method was compared to ordinary DQN spiking neural network approaches under the same environmental parameters. E-DQN realised the best performance with 73% and 74% less running time than SNN and DQN respectively.

In their prosposal, Wang et al. [98] suggested a novel UAV path planning technique based a layered priority experience replay DDQN algorithm (Layered PER-DDQN) in a complex battle field environment, where a priority label defines data to be extracted from the experience pool. The authors split the problem into two sub-problems layers. The first sub-problem was threat avoidance where the environment were modeled as a 2D graph. The second sub-problem was a collision avoidance where the environment were modeled as a 3D digital elevation map. A weighted action summation vector was used to sum the collision avoidance action vector and the threat avoidance action vector. The proposed Layered PER-DDQN approach was trained through 3000 episodes alongside with conventional DDQN proving a relatively quicker convergence and a higher average reward. In 250Km x 250Km x 2Km test environment, the proposed approach outperformed DDQN and A* in terms of path length and planning time.

Xie et al. in [99] introduced several improvements on DQN algorithm seeking for a better path planning solution for UAV in a large scale dynamic environment. A recurrent layer was added to the DQN network (DRQN) inorder to overcome blindness in early stage training. Furthermore, to speed up the algorithm, two approaches has been explored. An action selection strategy combining Q values and the reward (RQ) to attenuate meaningless exploration. An adaptive sampling (ADSA) process to distinguish important samples and normal ones based on rewards to rationalize UAV-environment interactions. Simulation was conducted in three different environments: static, dynamic and indoor. Five algorithms constructed from the prvious approaches has been compared: DRQ, ordinary DQN, RQ-DRQN, ADSA-DRQN and RQ-ADSA-DRQN. As a summary of the results, RQ-ADSA-DRQN had a relatively better global performance in the three scenarios regarding convergence speed and success rate indicators.

WANG et al. [100] suggested improvements to DQN algorithm in their proposal for a UAV autonomous navigation and collision avoidance problem in unknown environment relying on on board camera only. The authors first replaced original CNN network with faster R-CNN for image processing and obstacles identification. The second improvement was in the replay memory where the authors proposed a data deposit mechanism (DDM) that defines and sums three weighted classes of experiences: results experiences (RE) causing the end of an episode, danger experience (DE) related to obstacles detection in contrast to safety experience (SE) where UAV is away from obstacles. The obtained method was then a Faster-R-CNN-Data Deposit Mechanism-DQN (FR-DDM-DQN). Training of the FR-DDM-DQN algorithm was done through two parts where R-CNN and DDM-DQN algorithms were trained separately. FRDDM-DQN was tested in Unity3D simulation environment in three scenarios: static obstacles, dynamic obstacles and mixed environment. The FRDDM-DQN approach had a higher success rate in all scenarios outperforming methods without data deposit mechanism (DDM) such as FR-DDQN and FR-D3QN. It was noticeable also that faster R-CNN performed better than You Look Only Once (YOLO) algorithm since FR-DDM-DQN had a better success rate than YOLO-DDM-DQN in all scenarios.

Boulares et al. [101] developed a path planning algorithm for multiple UAV searching for floating targets in ocean environment. The idea started by applying a grid decomposition algorithm to the sea area and setting nodes and UAV was said to navigate between them. The nodes were set at the center of each grid, except centers outside the sea area whose nodes were kept nearest to borders inside that cell to prevent flight outside sea area. The sea area was then divided into sub-areas where a single UAV was assigned per sub-area. A simulation of target trajectory was settled based on Ekman Globcurrent data-set that takes into account wind and ocean current forces. Finally, a DQN algorithm was developed to for each UAV agent to simulated target search inside a sub-area. In simulation, a surface of 453 422 Km^2 was divided into seven 7 sub-areas where each agent is trained in 2000 episodes. The results has shown a high success rate between 96% and 100%. The search time and traveled distance were inversely proportional to the surface of the area.

Actor critic algorithms has an architecture constituted of two networks: an actor who selects the highest value action, a critic that evaluates that action [51] based on Temporal Difference (TD) [102, 103]. Han et al. [104] proposed an Experience-Shared Advantaged Actor Critic (ES-A2C) as a solution to multi-UAV path planning problem. It consisted of sharing prior exploration knowledge about the environment between UAV agents. This approach was tested in 3D dynamic simulated environment using: a single UAV, two UAVs and three UAVs. The results has shown that ES-A2C performs better that basic Actor Critic (AC) and Advantage Actor Critic (A2C) regarding average reward and convergence speed, especially in multiple-UAVs cases where it registered clearer dominance.

Jiménez et al. [105] conducted a comparative study between three algorithms: DQN, SARSA and A2C for UAV path planning. A virtual 3D neighborhood environment was designed in Microsoft Airsim with realistic obstacles where the UAV was said to navigate from a point A to a point B. The algorithm were then tested in 4000 episodes. The result shows that DQN algorithm was the only one among the three algorithms tor reach the target. A2C fell quickly into a local maximum and remained stuck there. The authors concluded that A2C algorithm has a faster convergence with a high data efficiency but requires a fine tuning to outperform DQN in complex environments especially.

Zhao et al. [106] introduced soft actor critic (SAC) algorithm with hindsight experience replay (HER) in a mobile robot path planning. The idea was to solve the problem of wasted experience in failed tasks by updating the target point for each sequence. This approach was tested in simulation through a 20x20 static environments. HER-SAC yields to a shorter path and a faster convergence to SAC and DDPG.

Zhou et al. [107] suggested several improvements to SAC algorithm in their solution to a UAV 3D online path planning problem. First, a self attention mechanism was inserted in to the actor network to face the large input data analysis processing and analysis. Furthermore, artificial gravitational and repulsive potential fields were injected into the reward mechanism so that to guide the UAV to the target and prevent collisions and hence accelerate convergence. The approach was trained and tested in a 3D 600x600x250 grid environment through static then dynamic scenarios. Simulation results in the static scenario has shown a better performance of the proposed enhanced SAC compared to ordinary SAC with 12.5% less planning time, 7.7% less steps to convergence and an overall improvement of 25.5%.

A strong generalization has been proven by the improved SAC in dynamic obstacles scenario with a maintained 100% average success rate.

Tian et al. [108] proposed a three step experience buffer deep deterministic policy gradient (TSEB-DDPG) to perform a fast path planning of a UAV in a 3D dynamic urban environment. First, the authors split the 3D environment into multi-layer 2D planes. Then, the agent was trained by Hybrid Learning Particle Swarm Optimization (HL-PSO) and the results were used as a prior knowledge for TSEB-DDPG. Finally, a three step sampling mechanism was adapted to basic DDPG algorithm experience buffer according to the learning stage: excellent transition buffer, collision transition buffer and fast transition buffer. Simulation results has proven that TSEB-DDPG outperforms classical DDPG and HL-PSO algorithms regarding planning time, path length and success rate.

HU et al. [109] also proposed an improved DDPG algorithm to plan UAV motion in complex unknown environment. The improvement introduced by the authors was the definition of Relevant Experience Learning (REL) in the experience pool. REL process started by experience pool splitting (EPS) to extract relevant experience according to danger degree measured by the distance to obstacles. Then Temporal Difference (TD-error) based priority experience replay (PER) was used to break correlation between adjacent similar experiences. And finally action selection was adjusted after experience learning so that REL experience is fully exploited. REP-DDPG approach was trained in 120 x 90 x 10 Km^3 static simulation environment through 5000 episodes alongside with Reward Classification (RC-DDPG), PER-DDPG, EPS-DDPG and classical DDPG algorithms. REL-DDPG had the fastest convergence in less than 3000 episodes and registered the highest average success rate of 88.92%. The approach was tested in three complex static scenarios and one complex dynamic scenario. UAV trained by REL-DDPG was the only one to reach the target in complex dynamic scenario.

Wang et al. [110] introduced a cumulative reward and an environment segmentation as improvements to DQN and DDPG algorithms to solve a UAV path planning problem. The proposed cumulative reward (CR) is computed inversely to the distance to the target and the obstacles' density in the neighborhood of the location. The region segmentation was based on a clustering process that divided the navigation area into partitions with eight borders where their connectivity was decisive in the reward function in order to avoid redundancy and local optima traps. Simulation results in 2D environment has shown that cumulative reward improved convergence speed of DQN by 30.8% and region segmentation enabled 99% local optima avoidance for DQN and 92% for DDPG.

Luo et al. [111] introduced an improved Twin Delay Temporal Difference Algorithm (I-TD3) to solve a UAV path planning problem. The improvement consisted of a prioritized sampling in the experience pool to identify important samples, and an average temporal difference (TD3) was adapted to overcome overestimation and avoid underestimation of Q values. This solution was tested in 4 different static 3D environments and 4 different dynamic 3D environments. The results after a 2000 episodes training has shown a higher performance of I-TD3 algorithm compared to SAC, DDPG and ordinary TD3 algorithm regarding registering the highest success and the lowest collision rate in all environments. The I-TD3 algorithm was also the best to resist to the increase of environment's complexity.

Fan et al. [112] suggested an algorithm based on interfered fluid dynamical system (IFDS) tuned by TD3 as a solution to an online UAV path planning problem. The authors established first the mathematical modeling of the UAV and the threat zones modeled as cylinders under dynamic constraints limiting flight angle and velocity. IFDS was designed as the basic method while TD3 generates the parameters of IFDS. Simulation results has shown that IFDS-TD3 avoided threat successfully in real time through a smooth path contrary to IFDS which fell into local optima. The main results issued from the reviewed reinforcement learning algorithms are summarized in Table III

V. Analysis and Discussion

A. Bibliometric Analysis

This review work focused on the recent contributions of machine learning in the solution of path planning problem for UAV. It covered 50 journal and conference articles. After the technical relevance of the paper, the source of the article was a crucial criterion in the selection. The reviewed works were issued from credible databases such as: IEEE, Elsevier, Springer Links and MDPI. The distribution of the reviewed journal papers and conference communications among databases is illustrated in Fig. 4.

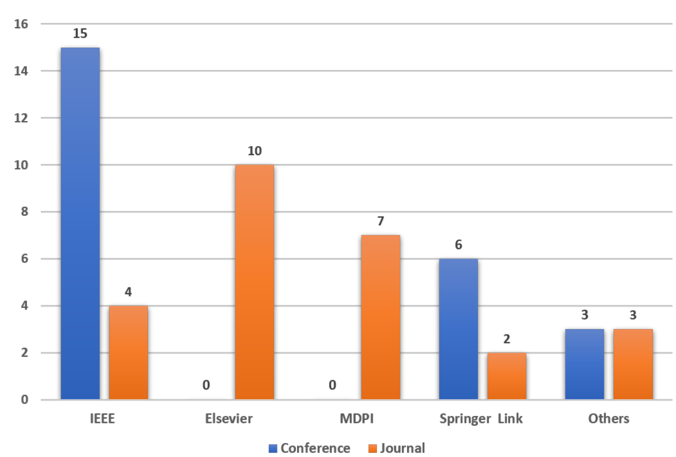

Fig. 4. Databases contribution in the review

TABLE III. Reinforcement Learning Proposals for UAV Path Planning

Algorithm	Author	Agent	Environment	Main Result	Year
RFQ	Gao et al. [86]	UAV	2D, static	faster convergence and a safer distance from obstacles than ordinary Q	2021
Q	Xie et al. [87]	UAV	2D, static	26% shorter path than ACO	2023
I-Q	Yan and Xiang[88]	UAV	2D, static	faster and more stable convergence than ordinary Q with relatively shorter path	2018
G-SARSA	Boming et al. [89]	UAV	2D, static	faster convergence with shorter path than Q and SARSA	2022
Dyna-Q	Huo et al. [90]	UAV	2D, static & dynamic	faster convergence	2022
2L-Q	Cui and Wang [91]	UAV	2D, static & dynamic	high success rate	2022
DQN	Anas et al.[93]	mobile robot	2D, static	higher success rate than Q and SARSA	2022
DDQN	Wang et al.[95]	UAV	2D, static	faster convergence than DQN	2022
D3QN	Yan et al.[96]	UAV	2D static & dynamic	higher success rate and better learning stability than DQN and DDQN	2022
E-DQN	Chao et al.[97]	UAV	3D, dynamic	less running time than SNN and DQN	2024
L-PER-DDQN	Wang et al.[98]	UAV	3D, static	shorter path lenght and completion time than DDQN and A*	2022
RQ-ASDA-DRQN	Xie et al.[99]	UAV	3D, static & dynamic	faster convergence and higher success rate than DQN, RQ-DQN, ADSA-DRQN and DRQ	2021
FR-DDM-DQN	WANG et al. [100]	UAV	3D, static & dynamic	higher success rate than FR-DDQN, FR-D3QN and YOLO-FDDM-DQN	2021
DQN	Boulares et al. [101]	miltiple UAV	2D, dynamic	high success rate	2021
ES-A2C	Han et al.[104]	UAVmultiple UAVs	3D dynamic	faster convergence and higher success rate compared to AC and A2C	2019
DQN	Jiménez et al.[105]	UAV	3D static	higher success rate than A2C which needs a fine tuning	2023
HER-SAC	Zhao et al.[106]	UAV	2D static	shorter path and faster convergence compared to DDPG and SAC	2023
I-SAC	Zhou et al.[107]	UAV	3D, static & dynamic	less completion time and faster convergence than original SAC	2023
TSEB-DDPG	Tian et al. [108]	multiple UAVs	3D, dynamic	better performance than original DDPG regarding planning time, path length and success rate	2023
REL-DDPG	HU et al. [109]	UAV	3D, static and dynamic	faster convergence and higher success rate than PER-DDPG and DDPG	2023
CR-RS-DQN	HU et al. [109]	UAV	2D, static	faster convergence than DQN	2023
I-TD3	Luo et al. [111]	UAV	3D dynamic	high resistance to complexity increase and higher success rate than SAC and DDPG	2024
IFDS-TD3	Fan et al. [112]	UAV	3D dynamic	better convergence and success rate in real time planning compared to IFDS alone	2020

The recency of the proposal was also an important criterion where the included papers were mostly dated from 2018. Fig. 5 illustrates the number of publications per year for each paradigm. Compared to other ML paradigms, there is an observable rising interest in reinforcement learning from researchers community for the design of path planning solutions.

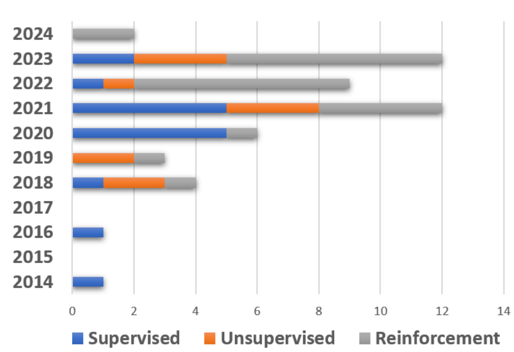

Fig. 5. Selected articles' years

B. Findings in Supervised Learning

Supervised learning was mainly present in this research through regression, SVM classifier and neural networks (Table I). Regression usefulness was mainly in path prediction and following relying on historical data and mean square error (MSE). Nevertheless, this classical approach struggles in high complexity environments with nonlinear variable interaction. Additionally, regression is a purely data-driven method whose outcome is strongly related to the quality of training. Hence, the model requires retraining to respond to dynamic changes leading to high computational cost with risks of overfitting.

SVM binary classifier gave excellent results in generating collision-free path. However, the high time-cost limits the scalability of this technique, which struggles in dynamic environments requiring continuously growing datasets.

Compared to previous techniques, neural networks and deep neural networks as particular supervised classification tools handles efficiently complex environments and large datasets with robustness against noise. DNN are used intensively in vision based navigation independently or within reinforcement techniques. Despite its relative high scalability and fast processing, DNN are hardware demanding to fit real-time requirements and large amount of training datasets.

The advantages and drawbacks of main SL techniques are summarized in Table IV.

TABLE IV. SL IN PATH PLANNING: ADVANTAGES AND DISADVANTAGES

SL Algorithm	Advantages	Drawbacks
Regression	Easy interpretation	Risk of overfitting Not suitable for complex environments
SVM	Efficient for safe navigation	High training cost Not suitable for high data sets
DNN	Efficient in complex environments	Data hungry and hardware demanding

C. Findings in Unsupervised Learning

Unsupervised learning algorithms contributed in path planning solutions by structuring the environment. Clustering algorithms strength was particularly noticed in multi-UAV and multi-target tasks (Table II). Clustering techniques like K-means were efficient in reducing flight time and hence saving resources by an optimized UAV task allocation. Despite their scalability to adapt clusters according to newer datasets, these algorithms struggles in dynamic cluster requirements where a frequent re-clustering leads to high computation cost and time lag in real-time applications.

The advantages and drawbacks of unsupervised clustering are summarized in Table V.

TABLE V. UL IN PATH PLANNING: ADVANTAGES AND DISADVANTAGES

SL Algorithm	Advantages	Drawbacks
Clustering	Efficient in resource management	Struggles with dynamic clusters

D. Findings in Reinforcement Learning

Reinforcement learning and deep reinforcement learning algorithms constitute the quantitatively dominant paradigm in this review. RL and DRL algorithms were proposed as solution indifferent scenarios (Table III). QL and SARSA were sufficient for simple environments where states are countable. These algorithms ensure fast convergence in such environments. Nevertheless, when state space get larger, memory requirements for storing a Q table become hard to satisfy.

Deep reinforcement learning uses deep neural networks approximations to deal with unlimited states. Deep Q network are scalable Q learning to deal with high dimensional state space. Additionally, actor critic algorithms are continuous action planners useful in smooth control navigation tasks.

Reinforcement learning and deep reinforcement learning algorithms minimizes the dependency to huge labeled datasets that supervised DNN required, by adopting reward based learning through the interaction with the environment. However most RL and DRL algorithms are highly sensitive to hyperparameters, like learning rate and neural networks batch size, which impact their learning and the balance of exploration (applying new actions) and exploitation (actions choice based on past experience).

The advantages and drawbacks of main UL techniques are summarized in Table VI.

TABLE VI. UL IN PATH PLANNING: ADVANTAGES AND DISADVANTAGES

RL Algorithm	Advantages	Drawbacks
Q	Probabilistic output Fast convergence at low cost	Limited scalability
DQN	Scalable to large state space	Higher computational cost
A2C	Continuous action generation	Higher computational cost

E. Enhancements and Hybridization

Following the previous analysis, some common features among machine learning paradigms are extracted and summarized in Table VII.

TABLE VII. INTER-PARADIGM COMPARISON

ML Paradigm	Advantages	Drawbacks
Supervised learning	Learning stability Efficient in static environments	Data dependency Limited adaptability
Unsupervised learning	Resource optimization with multi-UAVs	Not a standalone planning solution
Reinforcement learning	Less training time Adaptability do dynamic environments	Sensitivity to hyperparameters and reward shaping

To overcome limitations of machine learning approaches, authors proposed several improvement techniques:

In supervised learning, the combination of deep neural networks with heuristics and metaheuristic algorithms was conclusive. Based on its iterative approach, PSO can prevent DNN from falling into local minima and thus accelerating its convergence into optimal path. In addition, PSO can be used in DNN parameter and weights tuning by its capacity to high dimensional search and maintain multi-objective balance [64]. DNN itself can enhance traditional path search algorithm like A*, where it provides adaptive heuristics to A* in dynamic environments [65].

In most proposals, unsupervised learning constitute itself an enhancement to other planning algorithms. The addition of clustering algorithms like K-means improved resource management efficiency to many metaheuristic algorithms like ACO [113], GA [76] and SA [71]. The advantage of clustering was particularly sensed in WSN applications where UAV takes sensor range as cluster radius to minimize its flight path.

Deep reinforcement learning can be seen itself as an improvement to DNN where it reduced data dependency and training cost with an environment interaction. The proposed enhancements in DRL targeted mainly: convergence speed and learning stability by the mean of reward shaping. In many proposals in limited states environments, authors took into consideration the distance to the target in the Q function to assign higher rewards to closer states and hence accelerate convergence [89][88]. Dueling Double DQN (D3QN) introduced an evaluation network to reduce overestimation (DDQN [95]) and hence improve learning stability. In parallel, it evaluates state and action independently leading to more efficient learning[96]. Priority experience replay (PER) also attempts to accelerate convergence of DQN by giving priority to experiences with large TD error in target network update [95]. Hindsight experience replay was proposed to enhance learning efficiency of actor critic algorithms [106]. This technique consists of learning from failed episodes by setting new targets leading to a better exploration especially in sparse reward environments.

VI. Conclusion

UAV path planning has shown an increasing importance in recent years in parallel with the growing drones applications across diverse fields and important tasks. The autonomous data-driven nature of machine learning has made it a qualified solution for path planning problem. This review explores recent path planning proposals that used machine learning paradigms.

This review selected papers from credible databases and adopted paradigm-based taxonomy to classify them into: supervised, unsupervised and reinforcement learning. In light of proposals' results, an analysis has been built to highlight findings.

Supervised learning approaches has shown a good learning stability but with an extensive data dependency making them more suitable to static environments rather than dynamic scenarios. A hybrid combination with some metaheuristic algorithms like PSO could improve its real-time performance.

Unsupervised clustering algorithms were considered basically as optimizers rather than main planners. Clustering algorithms were found then as a hybrid combination with planners. These algorithms were efficient in resources optimization especially in muti-UAV tasks allocation, while it may struggle with dynamic clusters.

Reinforcement learning was the approach with a relatively better generalization to dynamic scenarios. Deep reinforcement learning approaches especially combines deep neural networks scalability with a reward based learning to get a better performance in real time. Nevertheless, their performance are highly sensitive to hyperparameters requiring a fine reward shaping to handle exploration-exploitation trade-off. Techniques like PER and HER were applied to enhance DQN convergence and SAC learning efficiency respectively. Improved versions of DQN like DDQN and D3QN were designed to enhance learning stability and convergence rate.

Despite the rigorous methodology adopted, there were some limitations of this paper that should be acknowledged. In many selected proposals, simulation results were not validated in real world experiment, this can be justified by regulatory restriction in some countries against drones. Although some authors resorted to high fidelity simulation environments that closely mimic real-world's, still an experimental evidence would strengthen the credibility of findings. Another weakness that could be mentioned is the absence of unified quantitative metrics for evaluation due to the adopted taxonomy which focus more on algorithm paradigm rather than the optimization criteria. This choice on the other hand conveys more to the objectives of the review.

In the end of this study, we are aware of some open challenges in machine learning based path planning whose solutions are actively under development:

- Reducing computational cost and thus hardware requirements for real-time applications.
- Leveraging multi-objective path planning.

In order to contribute to this development, the future of this research will be directed towards multi-objective path planning that balances fast real-time response with reducing hardware requirements.

List of Abbreviations

2L Two-layer.

ACO Ant colony optimization.
ADSA Adaptive sampling algorithm.

CHSA Cluster head selection algorithm.
CNN Convolutional neural network.
CSCA Coverage-based scanning clustering algorithm.

DDM Data deposit mechaniques.
DDPG Deep deterministic policy gradient.
DDQN Double deep Q network.
DTSPN Dubin travel sale's person with neighborhood.

E-DQN Event besed path planning.
ES-A2C Experience shared actor critic.

FNN Forward propagation neural network.

G-SARSA Guided state action reward state action.
GA Genetic algorithm.
GPR Gaussian process regression.

HALE High altitude long endurance.
HER Hindsight experience replay.

HTOL Horizontal take-off and landing.

I Improved.
IFDS Interfered fluid dynamical system.

KNN K neareast neighborhood.

L-PR Lasso regularized polynomial regression.
LASE Low attitude short endurance.
LASE Low attitude long endurance.
LR Linear regression.
LSTM Long short term memory.

MALE Medium altitude long endurance.
MLR Multiple linear regression.
MMAS Minimum Maximum.
MSVM Multiple support vector machine.

PCA Principal component analysis.
PER Priority experience replay.
PR Polynomial regression.
PSO Particle swarm optimization.

REL Relevant experience learning.
RF free.
RNN Recurrent neural network.

SAC Soft Actor Critic.
SAR Search and rescue.
SD Smart dust.
STCA Spatial-temporal clustering algorithm.
SVM Support vector machine.

TD3 Temporal difference.

UAV Unmanned aerial vehicle.

VTOL Vertical take-off and landing.

WRI Way refinement iteration.

REFERENCES

[1] J. Fan and M. A. Saadeghvaziri, "Applications of drones in infrastructures: Challenges and opportunities," *International Journal of Mechanical and Mechatronics Engineering*, vol. 13, no. 10, pp. 649–655, 2019, doi: 10.5281/zenodo.3566281.
[2] H. Shakhatreh *et al.*, "Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges," in *IEEE Access*, vol. 7, pp. 48572-48634, 2019, doi: 10.1109/ACCESS.2019.2909530.
[3] P. Mátyás and N. Máté, "Brief history of uav development," *Repüléstudományi Közlemények*, vol. 31, no. 1, pp. 155–166, 2019, doi: 10.32560/rk.2019.1.13.
[4] G. Muchiri and S. Kimathi, "A review of applications and potential applications of uav," in *Proceedings of the Sustainable Research and Innovation Conference*, pp. 280–283, 2022.
[5] A. Rejeb, A. Abdollahi, K. Rejeb, and H. Treiblmaier, "Drones in agriculture: A review and bibliometric analysis," *Computers and Electronics in Agriculture*, vol. 198, 2022, doi: 10.1016/j.compag.2022.107017.
[6] J. Fan and M. A. Saadeghvaziri, "Applications of drones in infrastructures: Challenges and opportunities," *International Journal of Mechanical and Mechatronics Engineering*, vol. 13, no. 10, pp. 649–655, 2019, doi: 10.5281/zenodo.3566281.
[7] S. M. S. Mohd Daud *et al.*, "Applications of drone in disaster management: A scoping review," *Science & Justice*, vol. 62, no. 1, pp. 30–42, 2022, doi: 10.1016/j.scijus.2021.11.002.
[8] A. Puente-Castro, D. Rivero, A. Pazos, and E. Fernandez-Blanco, "A review of artificial intelligence applied to path planning in UAV swarms," vol. 34, no. 1, pp. 153–170, 2022, doi: 10.1007/s00521-021-06569-4.
[9] A. Gasparetto, P. Boscariol, A. Lanzutti, and R. Vidoni, "Path planning and trajectory planning algorithms: A general overview," *Motion and Operation Planning of Robotic Systems: Background and Practical Approaches*, pp. 3–27, 2015, doi: 10.1007/978-3-319-14705-5_1.
[10] X. Li, J. Tupayachi, A. Sharmin, and M. Martinez Ferguson, "Drone-aided delivery methods, challenge, and the future: A methodological review," *Drones*, vol. 7, no. 3, 2023, doi: 10.3390/drones7030191.
[11] R. Masroor, M. Naeem, and W. Ejaz, "Resource management in uav-assisted wireless networks: An optimization perspective," *Ad Hoc Networks*, vol. 121, 2021, doi: 10.1016/j.adhoc.2021.102596.
[12] K. Okumura, F. Bonnet, Y. Tamura and X. Défago, "Offline Time-Independent Multiagent Path Planning," in *IEEE Transactions on Robotics*, vol. 39, no. 4, pp. 2720-2737, 2023, doi: 10.1109/TRO.2023.3258690.
[13] H. Sun, W. Zhang, R. Yu and Y. Zhang, "Motion Planning for Mobile Robots—Focusing on Deep Reinforcement Learning: A Systematic Review," in *IEEE Access*, vol. 9, pp. 69061-69081, 2021, doi: 10.1109/ACCESS.2021.3076530.
[14] M. Krichen, "Strengthening the security of smart contracts through the power of artificial intelligence," *Computers*, vol. 12, no. 5, 2023, doi: 10.3390/computers12050107.
[15] R. Choi, A. Coyner, J. Kalpathy-Cramer, M. Chiang, and J. Campbell, "Introduction to machine learning, neural networks, and deep learning," *Translational vision science & technology*, vol. 9, no. 2, 2020, doi: 10.1167/tvst.9.2.14.
[16] A. Al-Kaff, D. Martín, F. García, A. de la Escalera, and J. M. Armingol, "Survey of computer vision algorithms and applications for unmanned aerial vehicles," *Expert Systems with Applications*, vol. 92, pp. 447–463, 2018, doi: 10.1016/j.eswa.2017.09.033.
[17] V. S. Ajith and K. Jolly, "Unmanned aerial systems in search and rescue applications with their path planning: a review," *Journal of Physics: Conference Series*, vol. 2115, no. 1, 2021, doi: 10.1088/1742-6596/2115/1/012020.
[18] M. Hooshyar and Y.-M. Huang, "Meta-heuristic algorithms in uav path planning optimization: A systematic review (2018–2022)," *Drones*, vol. 7, no. 12, 2023, doi: 10.3390/drones7120687.
[19] H. S. Yahia and A. S. Mohammed, "Path planning optimization in unmanned aerial vehicles using meta-heuristic algorithms: a systematic review," *Environmental Monitoring and Assessment*, vol. 195, no. 30, 2023, doi: 10.1007/s10661-022-10590-y.
[20] A. Ait Saadi, A. Soukane, Y. Meraihi, A. Benmessaoud Gabis, S. Mirjalili, and A. Ramdane-Cherif, "UAV path planning using optimization approaches: A survey," *Archives of Computational Methods in Engineering*, vol. 29, pp. 4233–4284, 2022, doi: 10.1007/s11831-022-09742-7.
[21] Y. Zhang, W. Zhao, J. Wang, and Y. Yuan, "Recent progress, chal-

lenges and future prospects of applied deep reinforcement learning : A practical perspective in path planning," *Neurocomputing*, vol. 608, 2024, doi: 10.1016/j.neucom.2024.128423.

[22] M. Al-Shareeda, M. Ali, and S. Manickam, "Unmanned aerial vehicle: a review and future directions," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 2, pp. 778–786, 2023, doi: 10.11591/ijeecs.v30.i2.pp778-786.

[23] M. Akhloufi, S. Arola, and A. Bonnet, "Drones chasing drones: Reinforcement learning and deep search area proposal," *drone*, vol. 3, no. 3, 2019, doi: 10.3390/drones3030058.

[24] A. Otto, N. Agatz, J. Campbell, B. Golden, and E. Pesch, "Optimization approaches for civil applications of unmanned aerial vehicles (uavs) or aerial drones: A survey," *Networks*, vol. 72, no. 4, pp. 411–458, 2018, doi: 10.1002/net.21818.

[25] C. Lee, S. Kim, and B. Chu, "A survey: Flight mechanism and mechanical structure of the UAV," vol. 22, no. 4, pp. 719–743, 2021, doi: 10.1007/s12541-021-00489-y.

[26] M. Hassanalian and A. Abdelkefi, "Classifications, applications, and design challenges of drones: A review," *Progress in Aerospace Sciences*, vol. 91, pp. 99–131, 2017, doi: 10.1016/j.paerosci.2017.04.003.

[27] A. Watts, V. Ambrosia, and E. Hinkley, "Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use," *remote sensing*, vol. 4, no. 6, pp. 1671–1692, 2012, doi:10.3390/rs4061671.

[28] H.-y. Zhang, W.-m. Lin, and A.-x. Chen, "Path planning for the mobile robot: A review," *Symmetry*, vol. 10, no. 10, 2018, doi: 10.3390/sym10100450.

[29] M. Radmanesh, M. Kumar, P. H. Guentert, and M. Sarim, "Overview of path-planning and obstacle avoidance algorithms for uavs: A comparative study," *Unmanned systems*, vol. 6, no. 02, pp. 95–118, 2018, doi: 10.1142/S2301385018400022.

[30] O. Souissi, R. Benatitallah, D. Duvivier, A. Artiba, N. Belanger and P. Feyzeau, "Path planning: A 2013 survey," *Proceedings of 2013 International Conference on Industrial Engineering and Systems Management (IESM)*, p. 1-8, 2013.

[31] B. Han, T. Qu, X. Tong, J. Jiang, S. Zlatanova, H. Wang, and C. Cheng, "Grid-optimized uav indoor path planning algorithms in a complex environment," *International Journal of Applied Earth Observation and Geoinformation*, vol. 111, 2022, doi: 10.1016/j.jag.2022.102857.

[32] Y. U. Tursinovich, "Geometric modeling of threedimensional space and body," *Eurasian Journal of Physics,Chemistry and Mathematics*, vol. 5, no. 4, pp. 85–88, 2022.

[33] C. Zhang, H. Liu and Y. Tang, "Quantitative Evaluation of Voronoi Graph Search Algorithm in UAV Path Planning," *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*, pp. 563-567, 2018, doi: 10.1109/ICSESS.2018.8663950.

[34] M. Aleksandrov, S. Zlatanova, and D. J. Heslop, "Voxelisation algorithms and data structures: A review," *Sensors*, vol. 21, no. 24, 2021, doi: 10.3390/s21248241.

[35] X. Wu *et al.*, "A Non-Rigid Hierarchical Discrete Grid Structure and Its Application to UAVs Conflict Detection and Path Planning," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 6, pp. 5393-5411, 2022, doi: 10.1109/TAES.2022.3170323.

[36] J. Meyer, A. Sendobry, S. Kohlbrecher, U. Klingauf, and O. von Stryk, "Comprehensive simulation of quadrotor uavs using ros and gazebo," in *Simulation, Modeling, and Programming for Autonomous Robots*, vol. 7628, pp. 400–411, 2012, doi: 10.1007/978-3-642-34327-8_36.

[37] S. "Shah, D. Dey, C. Lovett, and A. Kapoor, "Airsim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and Service Roboticss*, vol. 5, pp. 621–635, 2018, doi: 10.1007/978-3-319-67361-5_40.

[38] M. Reda, A. Onsy, A. Y. Haikal, and A. Ghanbari, "Path planning algorithms in the autonomous driving system: A comprehensive review," *Robotics and Autonomous Systems*, vol. 174, 2024, doi: 10.1016/j.robot.2024.104630.

[39] Z. Tang and H. Ma, "An overview of path planning algorithms," in *IOP Conference Series: Earth and Environmental Science*, vol. 804, no. 2, 2024, doi: 10.1088/1755-1315/804/2/022024.

[40] J. S. Zelek and M. D. Levine, "Local-global concurrent path planning and execution," in *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 30, no. 6, pp. 865-870, 2000, doi: 10.1109/3468.895924.

[41] L. Liu, X. Wang, X. Yang, H. Liu, J. Li, and P. Wang, "Path planning techniques for mobile robots: Review and prospect," *Expert Systems with Applications*, vol. 227, 2023, doi: 10.1016/j.eswa.2023.120254.

[42] S. Qiu, J. Dai, and D. Zhao, "Path planning of an unmanned aerial vehicle based on a multi-strategy improved pelican optimization algorithm," *Biomimetics*, vol. 9, no. 10, 2024, doi: 10.3390/biomimetics9100647.

[43] Y. Shen, Y. Zhu, H. Kang, X. Sun, Q. Chen, and D. Wang, "Uav path planning based on multi-stage constraint optimization," *Drones*, vol. 5, no. 4, 2021, doi: 10.3390/drones5040144.

[44] G. Rebala, A. Ravi, and S. Churiwala, "Machine learning definition and basics," in *An Introduction to Machine Learning*, pp. 1–17, 2019, doi: 10.1007/978-3-030-15729-6_1.

[45] J. Alzubi, A. Nayyar, and A. Kumar, "Machine learning from theory to algorithms: An overview," *Journal of Physics: Conference Series*, vol. 1142, no. 1, 2018, doi: 10.1088/1742-6596/1142/1/012012.

[46] J. Peng, E. Jury, P. Dönnes, and C. Ciurtin, "Machine learning techniques for personalised medicine approaches in immune-mediated chronic inflammatory diseases: Applications and challenges," *Frontiers in Pharmacology*, vol. 12, 2021, doi: 10.3389/fphar.2021.720694.

[47] D. Sarkar, R. Bali, and T. Sharma, "Machine learning basics," in *Practical Machine Learning with Python: A Problem-Solver's Guide to Building Real-World Intelligent Systems*, pp. 3–65, 2018, doi: 10.1007/978-1-4842-3207-1_1.

[48] V. Nasteski, "An overview of the supervised machine learning methods," vol. 4, pp. 51–62, 2017, doi: 10.20544/HORIZONS.B.04.1.17.P05.

[49] P. Y C a, V. Pulabaigari, and B. E. Reddy, "Semi-supervised learning: a brief review," vol. 7, pp. 81–85, 2018, doi: 10.14419/ijet.v7i1.8.9977.

[50] S. Naeem, A. Ali, S. Anam, and M. Ahmed, "An unsupervised machine learning algorithms: Comprehensive review," *International Journal of Computing and Digital Systems*, vol. 13, no. 1, pp. 911–921, 2023, doi: 10.12785/ijcds/130172.

[51] F. AlMahamid and K. Grolinger, "Reinforcement Learning Algorithms: An Overview and Classification," *2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1-7, 2021, doi: 10.1109/CCECE53047.2021.9569056.

[52] S. Patil and S. Patil, "Linear with polynomial regression: Overview," *International Journal of Applied Research*, vol. 7, no. 8, pp. 273–275, 2021, doi: 10.22271/allresearch.2021.v7.i8d.8876.

[53] L. Meng, L. Yang, S. Ren, G. Tang, L. Zhang, F. Yang, and W. Yang, "An approach of linear regression-based UAV GPS spoofing detection," vol. 2021, pp. 1–16, 2021, doi: 10.1155/2021/5517500.

[54] P. Kumar, C. Sahu, D. Parhi, K. Pandey, and A. Chhotray,

"Static and dynamic path planning of humanoids using an advanced regression controller," *International Journal of Science & Technology*, vol. 26, no. 1, pp. 375–393, 2019, doi: 10.24200/sci.2018.5064.1071.

[55] H. Qu, K. Xing, and T. Alexander, "An improved genetic algorithm with co-evolutionary strategy for global path planning of multiple mobile robots," *Neurocomputing*, vol. 120, pp. 509–517, 2013, doi: 10.1016/j.neucom.2013.04.020.

[56] K. M. Koo *et al*., "A uav path planning method using polynomial regression for remote sensor data collection," in *Advances in Computer Science and Ubiquitous Computing*, vol. 536, pp. 428–433, 2020, doi: 10.1007/978-981-13-9341-9_74.

[57] M. Boulares and A. Barnawi, "A novel uav path planning algorithm to search for floating objects on the ocean surface based on object's trajectory prediction by regression," *Robotics and Autonomous Systems*, vol. 135, 2021, doi: 10.1016/j.robot.2020.103673.

[58] E. Yel and N. Bezzo, "GP-based Runtime Planning, Learning, and Recovery for Safe UAV Operations under Unforeseen Disturbances," *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2173-2180, 2020, doi: 10.1109/IROS45743.2020.9341641.

[59] M. A. Cano Lengua and E. A. Papa Quiroz, "A Systematic Literature Review on Support Vector Machines Applied to Classification," *2020 IEEE Engineering International Research Conference (EIRCON)*, pp. 1-4, 2020, doi: 10.1109/EIRCON51178.2020.9254028.

[60] M. Al-Naeem, M. M. Hafizur Rahman, A. Banerjee, and A. Sufian, "Support vector machine-based energy efficient management of uav locations for aerial monitoring of crops over large agriculture lands," *Sustainability*, vol. 15, no. 8, 2023, doi: 10.3390/su15086421.

[61] Y. Chen, W. Zu, G. Fan, and H. Chang, "Unmanned aircraft vehicle path planning based on svm algorithm," in *Foundations and Practical Applications of Cognitive Systems and Information Processing*, vol. 215, pp. 705–714, doi: 10.1007/978-3-642-37835-5_61.

[62] N. Morales, J. Toledo, and L. Acosta, "Path planning using a multiclass support vector machine," *Applied Soft Computing*, vol. 43, pp. 498–509, 2016, doi: 10.1016/j.asoc.2016.02.037.

[63] I. S. Asti, T. Agustinah and A. Santoso, "Obstacle Avoidance with Energy Efficiency and Distance Deviation Using KNN Algorithm for Quadcopter," *2020 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, pp. 285-291, 2020, doi: 10.1109/ISITIA49792.2020.9163788.

[64] A. Pandey, V. S. Panwar, M. E. Hasan, and D. R. Parhi, "V-REP-based navigation of automated wheeled robot between obstacles using PSO-tuned feedforward neural network," *Journal of Computational Design and Engineering*, vol. 7, no. 4, pp. 427–434, 2020, doi: 10.1093/jcde/qwaa035.

[65] G. Sanna, S. Godio and G. Guglieri, "Neural Network Based Algorithm for Multi-UAV Coverage Path Planning," *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1210-1217, 2021, doi: 10.1109/ICUAS51884.2021.9476864.

[66] Y. J. Choi, T. Rahim, I. N. A. Ramatryana and S. Y. Shin, "Improved CNN-Based Path Planning for Stairs Climbing in Autonomous UAV with LiDAR Sensor," *2021 International Conference on Electronics, Information, and Communication (ICEIC)*, pp. 1-7, 2021, doi: 10.1109/ICEIC51217.2021.9369805.

[67] Y. Liu, Z. Zheng, F. Qin, X. Zhang, and H. Yao, "A residual convolutional neural network based approach for real-time path planning," *Knowledge-Based Systems*, vol. 242, 2022, doi: 10.1016/j.knosys.2022.108400.

[68] X. Dai, Y. Mao, T. Huang, N. Qin, D. Huang, and Y. Li, "Automatic obstacle avoidance of quadrotor uav via cnn-based learning," *Neurocomputing*, vol. 402, pp. 346–358, 2020, doi: 10.1016/j.neucom.2020.04.020.

[69] D. Sartori, D. Zou, L. Pei and W. Yu, "CNN-based path planning on a map," *2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1331-1338, 2021, doi: 10.1109/ROBIO54168.2021.9739331.

[70] R. S. Nair and P. Supriya, "Robotic Path Planning Using Recurrent Neural Networks," *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1-5, 2020, doi: 10.1109/ICCCNT49239.2020.9225479.

[71] X. Yue and W. Zhang, "UAV Path Planning Based on K-Means Algorithm and Simulated Annealing Algorithm," *2018 37th Chinese Control Conference (CCC)*, pp. 2290-2295, 2018, doi: 10.23919/ChiCC.2018.8483993.

[72] J. Li, W. Li, and W. Zhang, "Path planning of UAV navigation mark inspection using a k-means clustering ACA," *Journal of Marine Science and Technology*, vol. 31, no. 3, 2023, doi: 10.51400/2709-6998.2705.

[73] S. Kim and J. Park, "Path planning with multiple UAVs considering the sensing range and improved k-means clustering in WSNs," *aerospace*, vol. 10, no. 11, 2023, doi: 10.3390/aerospace10110939.

[74] V. K. Chawra and G. P. Gupta, "Multiple UAV Path-Planning for Data Collection in Cluster-based Wireless Sensor Network," *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, pp. 194-198, 2020, doi: 10.1109/ICPC2T48082.2020.9071449.

[75] Y. Li, X. Meng, F. Ye, T. Jiang and Y. Li, "Path Planning Based on Clustering and Improved ACO in UAV-assisted Wireless Sensor Network," *2020 IEEE USNC-CNC-URSI North American Radio Science Meeting (Joint with AP-S Symposium)*, pp. 57-58, 2020, doi: 10.23919/USNC/URSI49741.2020.9321638.

[76] Y. Ma, H. Zhang, Y. Zhang, R. Gao, Z. Xu and J. Yang, "Coordinated Optimization Algorithm Combining GA with Cluster for Multi-UAVs to Multi-tasks Task Assignment and Path Planning," *2019 IEEE 15th International Conference on Control and Automation (ICCA)*, pp. 1026-1031, 2019, doi: 10.1109/ICCA.2019.8899987.

[77] P. Suseno and T. Wardana, "Unmanned air vehicle path planning for maritime surveillance using cluster-base method," *Aviation*, vol. 25, no. 3, pp. 211–219, 2021, doi: 10.3846/aviation.2021.14216.

[78] J. Chen, C. Du, Y. Zhang, P. Han and W. Wei, "A Clustering-Based Coverage Path Planning Method for Autonomous Heterogeneous UAVs," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 25546-25556, 2022, doi: 10.1109/TITS.2021.3066240.

[79] P. Xiao, N. Li, F. Xie, H. Ni, M. Zhang, and B. Wang, "Clustering-based multi-region coverage-path planning of heterogeneous uavs," *Drones*, vol. 7, no. 11, 2023, doi: 10.3390/drones7110664.

[80] J. Dai, Q. Hu, X. Liu, Y. Zhang, and J. Zhu, "Cluster head selection method of multiple uavs under covid-19 situation," *Computer Communications*, vol. 196, pp. 141–147, 2022, doi: 10.1016/j.comcom.2022.09.026.

[81] J. Faigl, P. Váňa, R. Pěnička, and M. Saska, "Unsupervised learning-based flexible framework for surveillance planning with aerial vehicles," *Journal of Field Robotics*, vol. 36, no. 1, pp. 270–301, 2019, doi: 10.1002/rob.21823.

[82] T. Kishimoto, H. Woo, R. Komatsu, Y. Tamura, H. Tomita, K. Shimazoe, A. Yamashita, and H. Asama, "Path planning for localization of radiation sources based on principal component analysis," *Applied Sciences*, vol. 11, no. 10, 2021, doi:

10.3390/app11104707.

[83] B. Ristic, M. Morelande, and A. Gunatilaka, "Information driven search for point sources of gamma radiation," *Signal Processing*, vol. 90, no. 4, pp. 1225–1239, 2010, doi: 10.1016/j.sigpro.2009.10.006.

[84] S. Kim and J. Park, "Path planning with multiple uavs considering the sensing range and improved k-means clustering in wsns," *Aerospace*, vol. 10, no. 11, 2023, doi: 10.3390/aerospace10110939.

[85] B. Jang, M. Kim, G. Harerimana and J. W. Kim, "Q-Learning Algorithms: A Comprehensive Classification and Applications," in *IEEE Access*, vol. 7, pp. 133653-133667, 2019, doi: 10.1109/ACCESS.2019.2941229.

[86] Y. Gao, Y. Li and Z. Guo, "A Q-learning based UAV Path Planning Method with Awareness of Risk Avoidance," *2021 China Automation Congress (CAC)*, pp. 669-673, 2021, doi: 10.1109/CAC53003.2021.9728342.

[87] X. Xie, T. Wang, Z. Zhu and S. Yang, "Q-Learning for Path Planning of a UAV under Energy Consumption Constraints," *2023 2nd International Symposium on Control Engineering and Robotics (ISCER)*, pp. 171-180, 2023, doi: 10.1109/ISCER58777.2023.00036.

[88] C. Yan and X. Xiang, "A Path Planning Algorithm for UAV Based on Improved Q-Learning," *2018 2nd International Conference on Robotics and Automation Sciences (ICRAS)*, pp. 1-5, 2018, doi: 10.1109/ICRAS.2018.8443226.

[89] H. Boming, L. Wei, M. Fuzeng and F. Huahao, "Research for UAV Path Planning Method Based on Guided Sarsa Algorithm," *2022 IEEE 2nd International Conference on Software Engineering and Artificial Intelligence (SEAI)*, pp. 220-224, 2022, doi: 10.1109/SEAI55746.2022.9832224.

[90] X. Huo, T. Zhang, Y. Wang, and W. Liu, "Dyna-q algorithm for path planning of quadrotor uavs," in *Methods and Applications for Modeling and Simulation of Complex Systems*, vol. 946, pp. 349–360, 2018, doi: 10.1007/978-981-13-2853-4_27.

[91] Z. Cui and Y. Wang, "UAV Path Planning Based on Multi-Layer Reinforcement Learning Technique," in *IEEE Access*, vol. 9, pp. 59486-59497, 2021, doi: 10.1109/ACCESS.2021.3073704.

[92] B. Jang, M. Kim, G. Harerimana and J. W. Kim, "Q-Learning Algorithms: A Comprehensive Classification and Applications," in *IEEE Access*, vol. 7, pp. 133653-133667, 2019, doi: 10.1109/ACCESS.2019.2941229.

[93] H. Anas, W. H. Ong, and O. A. Malik, "Comparison of deep q-learning, q-learning and sarsa reinforced learning for robot local navigation," in *Robot Intelligence Technology and Applications 6*, vol. 429, pp. 443–454, 2022, doi: 10.1007/978-3-030-97672-9_40.

[94] W. Luo, Q. Tang, C. Fu, P. Eberhard, Y. Shi, and Q. Tang, "Deep-sarsa based multi-uav path planning and obstacle avoidance in a dynamic environment," in *Advances in Swarm Intelligence*, 2018, doi: 10.1007/978-3-319-93818-910.

[95] Y. Wang, C. Jiang, T. Ren, Z. Yin, L. Liu, L. Jiang, G. Gu, X. Wu, and W. Ren, "Uav path planning based on ddqn for mountain rescue," in *Intelligent Robotics and Applications*, vol. 13458, pp. 509–516, 2022, doi: 10.1007/978-3-031-13841-6_46.

[96] C. Yan, X. Xiang, and C. Wang, "Towards real-time path planning through deep reinforcement learning for a UAV in dynamic environments," *Journal of Intelligent & Robotic Systems*, vol. 98, no. 2, pp. 297–309, 2020, doi: 10.1007/s10846-019-01073-3.

[97] Y. Chao, R. Dillmann, A. Roennau, and Z. Xiong, "E-dqn-based path planning method for drones in airsim simulator under unknown environment," *Biomimetics*, vol. 9, no. 4, 2024, doi: 10.3390/biomimetics9040238.

[98] W. Wang *et al.*, "A novel uav path planning method based on layered per-ddqn," in *The Proceedings of the 2021 Asia-Pacific International Symposium on Aerospace Technology (APISAT 2021), Volume 2*, vol. 913, pp. 693–702, 2023, doi: 10.1007/978-981-19-2635-8_51.

[99] R. Xie, Z. Meng, L. Wang, H. Li, K. Wang and Z. Wu, "Unmanned Aerial Vehicle Path Planning Algorithm Based on Deep Reinforcement Learning in Large-Scale and Dynamic Environments," in *IEEE Access*, vol. 9, pp. 24884-24900, 2021, doi: 10.1109/ACCESS.2021.3057485.

[100] F. Wang, X. Zhu, Z. Zhou, and Y. Tang, "Deep-reinforcement-learning-based uav autonomous navigation and collision avoidance in unknown environments," *Chinese Journal of Aeronautics*, vol. 37, no. 3, pp. 237–257, 2024, doi: 10.1016/j.cja.2023.09.033.

[101] M. Boulares, A. Fehri, and M. Jemni, "Uav path planning algorithm based on deep q-learning to search for a floating lost target in the ocean," *Robotics and Autonomous Systems*, vol. 179, 2024, doi: 10.1016/j.robot.2024.104730.

[102] A. K. Shakya, G. Pillai, and S. Chakrabarty, "Reinforcement learning algorithms: A brief survey," *Expert Systems with Applications*, vol. 231, 2023, doi: 10.1016/j.eswa.2023.120495.

[103] H. Kumar, A. Koppel, and A. Ribeiro, "On the sample complexity of actor-critic method for reinforcement learning with function approximation," vol. 112, no. 7, pp. 2433–2467, 2023, doi: 10.1007/s10994-023-06303-2.

[104] X. Han, J. Wang, Q. Zhang, X. Qin and M. Sun, "Multi-UAV Automatic Dynamic Obstacle Avoidance with Experience-shared A2C," *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 330-335, 2019, doi: 10.1109/WiMOB.2019.8923344.

[105] G. A. Jiménez, A. de la Escalera Hueso, and M. J. Gómez-Silva, "Reinforcement learning algorithms for autonomous mission accomplishment by unmanned aerial vehicles: A comparative view with dqn, sarsa and a2c," *Sensors*, vol. 23, no. 21, 2023, doi: 10.3390/s23219013.

[106] T. Zhao, M. Wang, Q. Zhao, X. Zheng, and H. Gao, "A path-planning method based on improved soft actor-critic algorithm for mobile robots," *Biomimetics*, vol. 8, no. 6, 2023, doi: 10.3390/biomimetics8060481.

[107] Y. Zhou, J. Shu, H. Hao, H. Song, and X. Lai, "UAV 3d online track planning based on improved SAC algorithm," vol. 46, no. 12, 2024, doi: 10.1007/s40430-023-04570-7.

[108] S. Tian, Y. Li, X. Zhang, L. Zheng, L. Cheng, W. She, and W. Xie, "Fast uav path planning in urban environments based on three-step experience buffer sampling ddpg," *Digital Communications and Networks*, vol. 10, no. 4, pp. 813–826, 2023, doi: 10.1016/j.dcan.2023.02.016.

[109] Z. Hu, X. Gao, K. Wan, Y. Zhai, and Q. Wang, "Relevant experience learning: A deep reinforcement learning method for uav autonomous motion planning in complex unknown environments," *Chinese Journal of Aeronautics*, vol. 34, no. 12, pp. 187–204, 2021, doi: 10.1016/j.cja.2020.12.027.

[110] Z. Wang, S. X. Ng and M. EI-Hajjar, "Deep Reinforcement Learning Assisted UAV Path Planning Relying on Cumulative Reward Mode and Region Segmentation," in *IEEE Open Journal of Vehicular Technology*, vol. 5, pp. 737-751, 2024, doi: 10.1109/OJVT.2024.3402129.

[111] X. Luo, Q. Wang, H. Gong and C. Tang, "UAV Path Planning Based on the Average TD3 Algorithm With Prioritized Experience Replay," in *IEEE Access*, vol. 12, pp. 38017-38029, 2024, doi: 10.1109/ACCESS.2024.3375083.

[112] J. Fan, Z. Wang, J. Ren, Y. Lu and Y. Liu, "UAV online path planning technology based on deep reinforcement learning," *2020 Chinese Automation Congress (CAC)*, pp. 5382-5386, 2020, doi:

10.1109/CAC51589.2020.9327752.
[113] J. Li, W. Li, and W. Zhang, "Path planning of uav navigation mark inspection using a k-means clustering aca," *Journal of Marine Science and Technology*, vol. 31, no. 3, 2023, doi: 10.51400/2709-6998.2705.