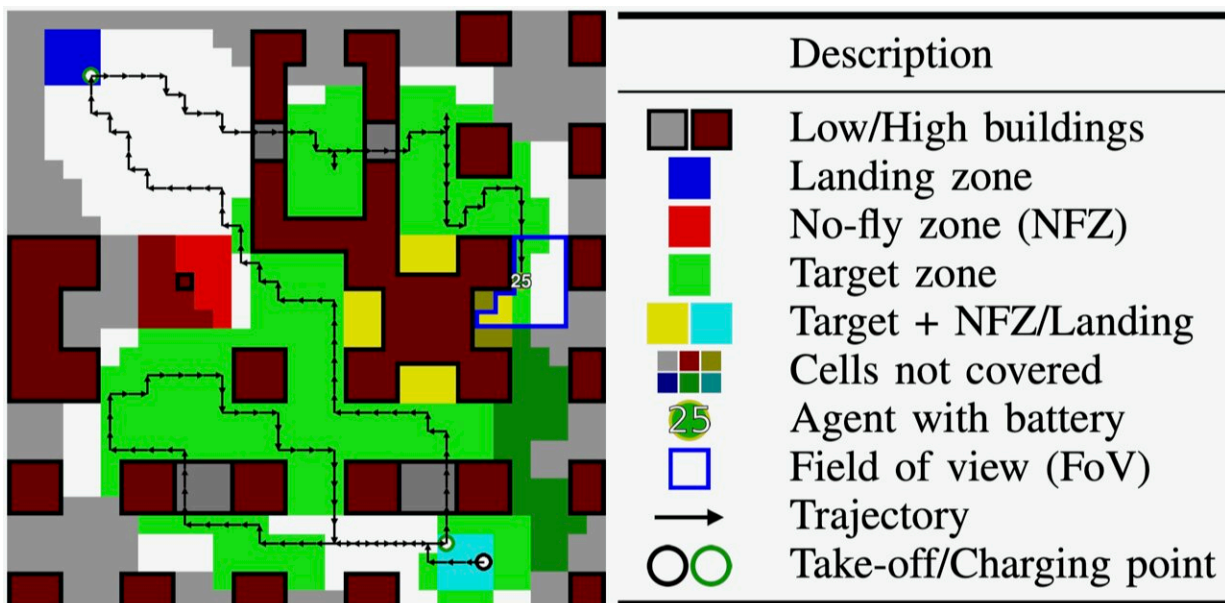# A reinforcement learning-based method to plan the coverage path and recharging of unmanned aerial vehicles

September 26 2023, by Ingrid Fadelli

The coverage path planning scenario in which the agent needs to find a trajectory that covers all of the green areas with its field of view without flying into obstacles and with recharging in blue landing zones. Credit: Theile et al.

Unmanned aerial vehicles (UAVs), commonly known as drones, have already proved invaluable for tackling a wide range of real-world problems. For instance, they can assist humans with deliveries, environmental monitoring, film-making and search & rescue missions.

While the performance of UAVs improved considerably over the past decade or so, many of them still have relatively short battery lives, thus they can run out of power and stop operating before completing a mission. Many recent studies in the field of robotics have thus been aimed at improving these system's battery life, while also developing computational techniques that allow them to tackle missions and plan their routes as efficiently as possible.

Researchers at Technical University of Munich (TUM) and University of California Berkeley (UC Berkeley) have been trying to devise better solutions for tackling the commonly underlying research problem, which is known as coverage path planning (CPP). In a recent paper pre-published on *arXiv*, they introduced a new reinforcement learning-based tool that optimizes the trajectories of UAVs throughout an entire mission, including visits to charging stations when their battery is running low.

"The roots of this research date back to 2016, when we started our research on "solar-powered, long-endurance UAVs," Marco Caccamo, one of the researchers who carried out the study, told Tech Xplore.

"Years after the start of this research, it became clear that CPP is a key component to enabling UAV deployment to several application domains like digital agriculture, search and rescue missions, surveillance, and many others. It is a complex problem to solve as many factors need to be considered, including collision avoidance, camera field of view, and battery life. This motivated us to investigate reinforcement learning as a potential solution to incorporate all these factors."

In their previous works Caccamo and his colleagues tried to tackle simpler versions of the CPP problem using reinforcement learning. Specifically, they considered a scenario in which a UAV had battery constraints and had to tackle a mission within a limited amount of time

(i.e., before its battery run out).

In this scenario, the researchers used reinforcement learning to allow the UAV to complete as much of a mission or move through as much space as possible with a single battery charge. In other words, the robot could not interrupt the mission to recharge its battery, subsequently re-starting from where it stopped before.

"Additionally, the agent had to learn the safety constraints, i.e., collision avoidance and battery limits, which yielded safe trajectories most of the time but not every time," Alberto Sangiovanni-Vincentelli explained. "In our new paper, we wanted to extend the CPP problem by allowing the agent to recharge so that the UAVs considered in this model could cover a much larger space. Furthermore, we wanted to guarantee that the agent does not violate safety constraints, an obvious requirement in a real-world scenario. "

A key advantage of reinforcement learning approaches is that tend to generalize well across different cases and situations. This means that after training with reinforcement learning methods, models can often tackle problems and scenarios that they did not encounter before.
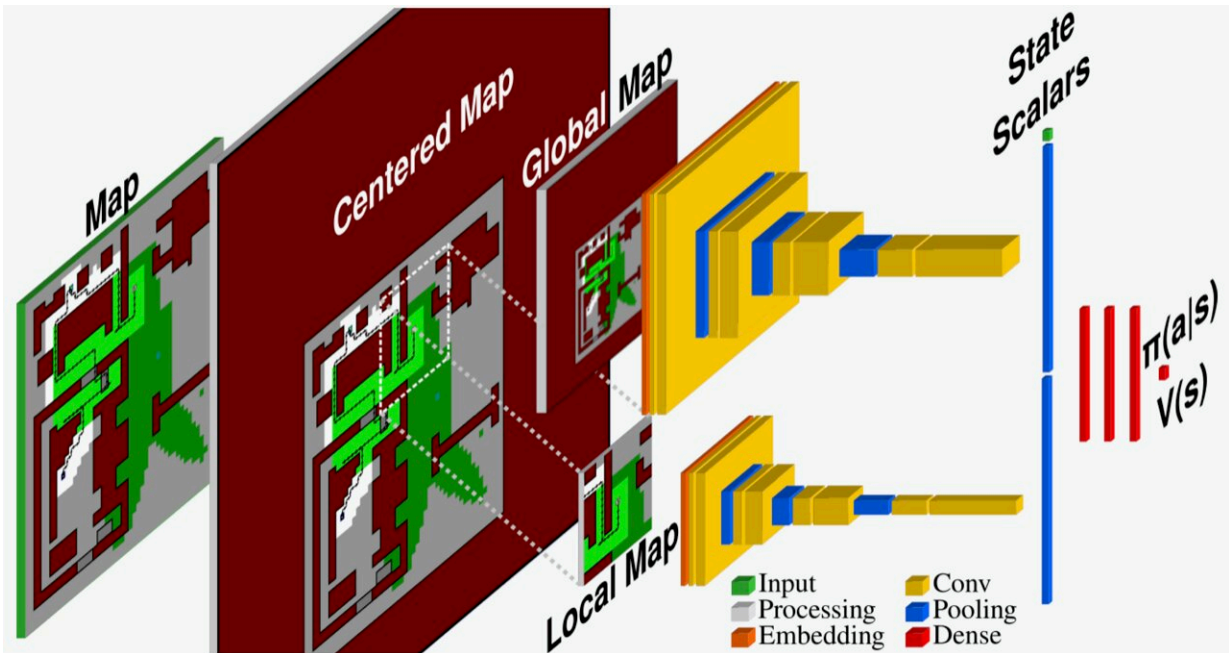
How the reinforcement learning agent observes and processes the environment. It observes a map representation of the problem, which it then centers around its position. It then compresses the entire centered map into a global map with reduced resolution and a full-resolution local map, which shows only the immediate vicinity of the agent. The agent then processes these two maps through different neural network layers until the agent outputs an action decision. Credit: Theile et al.

This ability to generalize greatly depends on how a problem is presented to the model. Specifically, the deep learning model should be able to look at the situation at hand in a structured way, for instance in the form of a map.

To tackle the new CPP scenario considered in their paper, Caccamo, Sangiovanni-Vincentelli and their colleagues developed a new reinforcement learning-based model. This model essentially observes and processes the environment in which a UAV is moving, which is

represented as a map, and centers it around its position.

Subsequently, the model compresses the entire 'centered map' into a global map with lower resolution and a full-resolution local map showing only the robot's immediate vicinity. These two maps are then analyzed to optimize trajectories for the UAV and decide its future actions.

"Through our unique map processing pipeline, the agent is able to extract the information it needs to solve the coverage problem for unseen scenarios," Mirco Theile said. "Furthermore, to guarantee that the agent does not violate the safety constraints, we defined a safety model that determines which of the possible actions are safe and which are not. Through an action masking approach, we leverage this safety model by defining a set of safe actions in every situation the agent encounters and letting the agent choose the best action among the safe ones."

The researchers evaluated their new optimization tool in a series of initial tests and found that it significantly outperformed a baseline trajectory planning method. Notably, their model generalized well across different target zones and known maps, and could also tackle some scenarios with unseen maps.

"The CPP problem with recharge is significantly more challenging than the one without recharge, as it extends over a much longer time horizon," Theile said. "The agent needs to make long-term planning decisions, for instance deciding which target zones it should cover now and which ones it can cover when returning to recharge. We show that an agent with map-based observations, safety model-based action masking, and additional factors, such as discount factor scheduling and position history, can make strong long-horizon decisions."

The new reinforcement learning-based approach introduced by this team of research guarantees the safety of a UAV during operation, as it only

allows the agent to select safe trajectories and actions. Concurrently, it could improve the ability of UAVs to effectively complete missions, optimizing their trajectories to points of interest, target locations and charging stations when their battery is low.

This recent study could inspire the development of similar methods to tackle CPP-related problems. The team's code and software [is publicly available on GitHub](#), thus other teams worldwide could soon implement and test it on their UAVs.

"This paper and our previous work solved the CPP problem in a discrete grid world," Theile added. "For future work, to get closer to real-world applications, we will investigate how to bring the crucial elements, map-based observations and safety action masking into the continuous world. Solving the problem in continuous space will enable its deployment in real-world missions such as smart farming or [environmental monitoring](#), which we hope can have a great impact."

More information: Mirco Theile et al, Learning to Recharge: UAV Coverage Path Planning through Deep Reinforcement Learning, *arXiv* (2023). [DOI: 10.48550/arxiv.2309.03157](#)