



Neuro-Fuzzy Probabilistic Control with Deep Q-Learning for Robust Multi-Robot Path Planning under Dynamic Obstacles

Dr. K.VIJAY KUMAR ¹, Dr.K.KANTHA RAO ²

kvsagar1987@gmail.com ¹, kkrao.konka@gmail.com ²

1. Associate Professor, Department of Mechanical Engineering, MallaReddy (MR) Deemed to be university, Kompally, Medchal, T.S, INDIA
2. Professor, Department of Mechanical Engineering, MallaReddy (MR) Deemed to be university, Kompally, Medchal, T.S, INDIA

How to Cite this Article:

KUMAR, K. (2026). Neuro-Fuzzy Probabilistic Control with Deep Q-Learning for Robust Multi-Robot Path Planning under Dynamic Obstacles. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(<i>02</i>).
<https://doi.org/10.55041/ijcope.v2i2.002>

License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i2.002>

ABSTRACT

The method consists of two fuzzy level controllers architecture based on a fuzzy probabilistic control and an Adaptive Neuro-Fuzzy Inference System (ANFIS). Each robot has low level probabilistic fuzzy controller to eliminate the stochastic uncertainties as well as to make the multi-robots team navigates from the start point to the target point without any dangerous collision. In addition, the robot collaboration scheme is highly depends on the conditions of each robot, such as its position and velocity. However, the conventional method does not actively cope with variable situations since each robot has difficulty to recognize the moving robot around it as an obstacle or a cooperative robot. To compensate for these shortcomings, we apply Deep q learning to strengthen the learning algorithm combined with CNN algorithm, which is needed to analyze the situation efficiently. This article aims to bring a brief review of the state-of-the-art NNs for the complex nonlinear systems by summarizing recent progress of NNs in both theory and practical applications. Specifically, this survey also reviews a number of NN based robot control algorithms, including NN based manipulator control, NN based human-robot interaction, and NN based cognitive control. Especially, the history and applications of numerous heuristic methods in MP is investigated. Simultaneously, a global backtracking mechanism guides the robot to move to the next unvisited area quickly, taking the use of the explored global environmental information. What's more, the authors extend their CCPP algorithm to a multi-robot system with a market-based bidding

process which could deploy the coverage time. Experiments of apartment-like scenes show that the authors' proposed algorithm can guarantee an efficient collision-free coverage in dynamic environments. The proposed method performs better than related approaches on coverage rate and overlap length.

Keywords: MRS, Path planning, Adaptive Neuro-Fuzzy Inference System (ANFIS), Deep Q-Network (DQN) Convolutional Neural Networks



Convolutional Neural Networks

1. INTRODUCTION

The use of cooperative robots system dramatically increases since it is appropriate solution in relation to performance, efficiency and reliability. The development of such cooperation system is one of the most demanding goals in artificial intelligence and robotics researches. The first study about multi-robot systems concerns certainly the reconfigurable robots with the project Cellular Robotics System which is a decentralized, hierarchical architecture (Fukuda & Nakagawa, 1987; Fukuda & Iritani, 1995). The swarm robot system where the multi-robot system is composed of a large number of homogeneous robots aims to represent a group of simple robots, self-organizing in new patterns. In addition, the control strategy for this system is based on individual behavior where each robot can perceive other robots by using proximity sensors (Beni, 1988). The first work on coordination among multiple agents was inspired by the study of biological systems for coordinating the animal motions (Parker, 1993). This work motivated significant efforts in the study of multi-robots formation (Wang, 1991; Seanor, 2006).

In past decades, the NN technique has been studied Extensively in areas such as control engineering, aerospace, medicine, automotive, psychology, economics, energy science, and many other fields [4–7]. It has been reported that NN can approximate any unknown continuous nonlinear function by overlapping the outputs of each neuron. Moreover, the approximation errors could be made arbitrarily small by choosing sufficient neurons. This enables us to deal with control problems for complex nonlinear systems [8]. In addition to system modeling and control, NN has also been successfully applied in various fields such as pattern recognition and signal processing. And NN has been extensively used for functions approximation, such as to compensate for the effect of unknown dynamics in nonlinear systems. The NN control has been proved to be effective for controlling uncertain nonlinear systems and demonstrated superiority in many aspects.

The robot learns the next action based on the learned data when it selects the next action, and after several learning, it moves to the closest target. By interacting with the environment, robots exhibit new and complex behaviors rather than existing behaviors. The existing analytical methods suffer from

adaptation to complex and dynamic systems and environments. By using Deep q learning and CNN reinforcement learning is performed on the basis of image, and the same data as the actual multi-robot is used to compare it with the existing algorithms.

In the proposed algorithm, the global image information in the multi-robot provides the with higher autonomy comparing with conventional robots. In this paper, we propose a noble method for a robot to move quickly to the target point by using reinforcement learning for path planning of a multi-robot system. In this case, reinforcement learning is a Deep q learning that can be used in a real mobile robot environment by sharing q parameters for each robot. In various 2D environments such as static and dynamic environment, the proposed algorithm spends less searching time than other path planning algorithms.

Great efforts have been made on CCPP methods in known and stationary environments. However, few methods pay attention to dynamic environments with non-stationary obstacles. Luo and Yang proposed a bio-inspired method where the dynamics of each position on the map is topologically organized in a network. However, in the Luo's method, the robot is easily trapped in a situation named deadlock where all the neighbouring locations are neither obstacles nor visited locations. The network inspired by the biologic shunting model is lack of global information to escape from deadlocks quickly. Both the single-robot based approaches and the approaches using multiple robots have been developed, such as Multi-Robot Spanning Tree Coverage (MSTC) Multi-Robot Forest Backtracking Spiral Algorithm Cooperative Multi-robot and Boustrophedon and Backtracking mechanism. These approaches reduce the coverage time in general.

2. HEURISTIC METHODS

The abovementioned classic approaches suffer from many drawbacks, such as high time complexity in high dimensions, and trapping in local minima, which makes them inefficient in practice. In order to improve the efficiency of Classic methods, Probabilistic algorithms have been developed, including Probabilistic Roadmaps (PRM) and Rapidly exploring Random Trees (RRT), with major advantages is high-speed implementation. Also other approaches exist in RMP such as Level set and Linguistic Geometry (LG). To fix the local minima



problem, many **Heuristic** and Meta-heuristic algorithms are used in RMP. For example, a combination of the **Simulated Annealing (SA)** technique and PF remedies this problem. Other approaches discussed in this paper include Artificial Neural Network (ANN), Genetic Algorithms (GA), Particle Swarm Optimization (PSO), Ant Colony (ACO), Stigmergy, Wavelet Theory, Fuzzy Logic (FL) and Tabu Search (TS). Heuristic algorithms do not guarantee to find a solution, but if they do, are likely to do so much faster than deterministic methods.

A. Neural Networks (NN)

The concept of using Neural Networks for RMP was first used. A novel biologically-inspired general neural network approach exists for real-time collision-free RMP in a dynamic environment. This general model can be applied to point mobile robots, manipulator robots, car-like robots, and multi-robot systems. The state space of the NN is the configuration space of the robot, and the dynamically varying environment is represented by the dynamic activity landscape of the neural network. The target globally attracts the robot in whole state space, while the obstacles locally push the robot away to avoid collisions. The real-time robot motion is planned through the dynamic activity landscape of the neural network without explicitly searching over the free space or the collision-free paths, without explicitly optimizing any cost function, without any prior knowledge of the dynamic environment, without any learning process, and without any local collision checking procedures. Therefore, the model algorithm is computationally efficient. In an NN approach to path planning for two dimensional robot motion is developed. Also in a neural network approach for the local navigation of a mobile robot via Perception maps is presented. In 1995, the collision identification between convex polyhedra using neural networks is implemented. A cache-genetic-based modular fuzzy neural network is presented for robot path planning. Front zek Basis Function networks. An NN model is developed in to real-time MP and control of robot manipulators. RMP problem is solved in using Hopfield neural networks in a fuzzified environment. Also in 2003, a Non-learning ANN approach to MP for the Pioneer robot is extended. An NN approach is presented in for dynamic task assignment of multi robots. Eventually, RL-ART2 NN-based mobile robot path planning is developed in 2007.

B. Genetic Algorithms (GA) The idea of using Genetic algorithms for RMP was first used. The application of GA to the mobile robot path planning problem requires the development of a suitable ‘chromosome’ representation of the path, a path guidance mechanism, a method to cater for obstacle avoidance, and an appropriate constraint definition providing mechanisms to minimize path distance, as well as providing smooth paths. It is assumed that the environment is static and known. For the purpose of representation the working area is discretized in bit-map form with the obstacles providing coloring to aid the obstacle avoidance. In a genetic approach for generation of collision-free paths is presented. In an approach on the basis of GA for planning multi-paths is presented. In 1997, another approach from GA for solving shortest path problem is used. A genetic-fuzzy algorithm is introduced in for mobile robot navigation among static obstacles. A multiple path planning for a group of mobile robots in a 2D environment using GA is developed. Zein-Sabatto constructed a multiple path planning for a group of mobile robots in a 3D environment using GA. A novel GA searching approach for dynamic constrained multicast routing is developed. Also in a Parallel GA is used for search and constrained multi objective optimization. An optimum path planning for mobile robots based on a Hybrid GA has been extended. Finally, A hybrid method for shared-path protections in WDM Networks under SRLG constraints has been developed.

C. Particle Swarm Optimization (PSO) Particle Swarm Optimization (PSO) was invented by Kennedy and Eberhart in 1995 while attempting to simulate the choreographed, graceful motion of swarms of birds as part of a sociocognitive study investigating the notion of “collective intelligence” in biological populations. In PSO, a set of randomly generated solutions (initial swarm) propagates in the design space towards the optimal solution over a number of iterations (moves) based on large amount of information about the design space that is assimilated and shared by all members of the swarm. PSO is inspired by the ability of flocks of birds, schools of fish, and herds of animals to adapt to their environment, find rich sources of food, and avoid predators by implementing an “information sharing” approach, hence, developing an evolutionary advantage. An algorithm for path planning for mobile robot using PSO with mutation operator is developed. In 2005, an approach for obstacle avoidance with



multi-objective optimization by PSO in dynamic environment is presented. Also, an algorithm is developed for robot path planning using PSO of Ferguson Splines. Obstacle avoidance path planning for soccer robots using PSO has been extended. Finally, a smooth path planning of a mobile robot using Stochastic PSO is implemented.

3. PROPOSED METHOD

The framework of the proposed approach is illustrated in Fig. 1 with a single robot case as an example. The prior knowledge of the environment is completely unknown, by gathering the local information via sensors. The robot is expected to cover all accessible grids. The first two subsections describe the proposed CCP algorithm in single robot case. In the bio-inspired motion coverage process, the robot covers unvisited regions by calculating the activities of the neuron network; the network was constructed by the short-memory CCP method and with a prior template to reduce deadlocks. When a deadlock situation occurs, the backtracking process activates, this process including three steps:

- (i) searching for candidate backtracking points and updating the backtracking list; (ii) selecting the best backtracking point; (iii) planning the shortest path from a deadlock point to a backtracking point. The coverage process will not finish until all points have

visited. What's more, in Section 4, we extend the approach to the multi-robot case with a market-bidding process. Overall, the analysis of the approach in a dynamic environment is given in the last subsection.

3.1 Shunting short-memory-based coverage path planning with a prior template

A human brain could use the short-memory model to deal with the information in a dynamic environment. We adapt the model to path planning problems, especially in the coverage task.

3.1.1 Landscape of neural activities:

The core of the algorithm is to propose a neural network for the coverage task.

The neural activities could represent the coverage state of the robot. Through the neural activity propagation, the robot will be attracted by the activities of neighbour neurons, in the way that inspired by shunting short-term memory mechanism. A short-memory model is introduced by Hodgkin and Huxley. It describes how information to transport between the paths of membranes in a biological neural system of humans. This model understands the real-time adaptive behaviour of individuals to complex and dynamic environmental contingencies. In the neural activities across the membranes are described as follows:

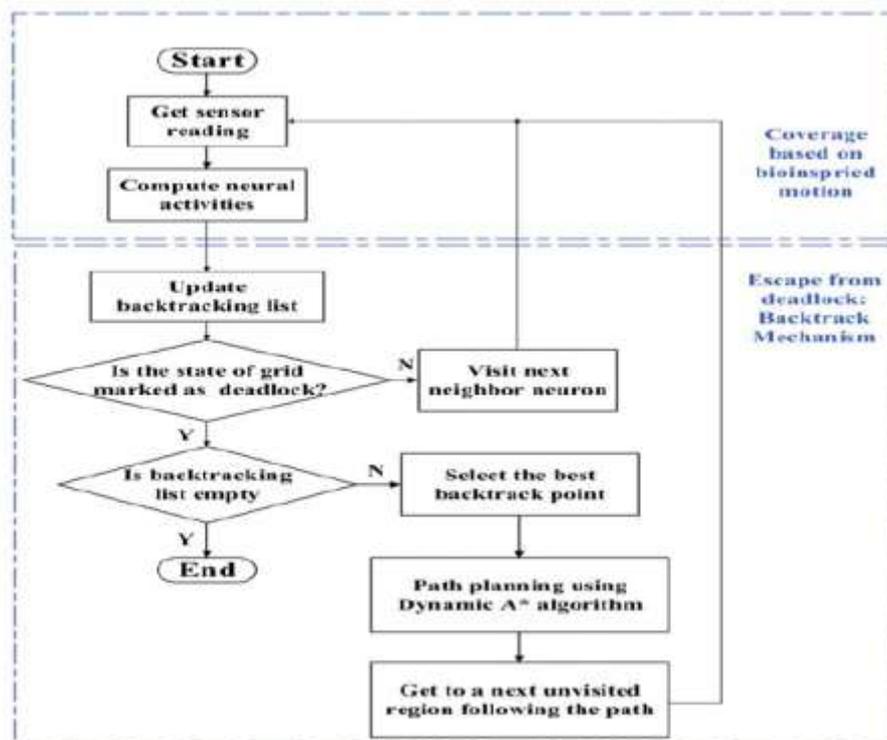


Fig. 1 Flowchart of the proposed method



$$\frac{dx_i}{dt} = -Ax_i + (B - x_i) \left([I_i]^+ + \sum_{j=1}^k \omega_{ij} [x_j]^+ \right) - (D + x_i) [I_i]^- \quad (1)$$

Parameters A, B and D are non-negative constants representing the passive decay rate, and the upper and lower bounds of the neural activity, respectively. k is the number of neural connections from the ith neuron to its neighbours.

3.2 Global backtracking mechanism

We active a global tracking mechanism to escape from the deadlock situation quickly. Our backtracking mechanism mainly lies in two aspects: First, while updating the BTlist, a restriction is imposed according to the spatial characters. Second, a greedy criterion is used to select the best backtracking point, and then robot escapes from deadlocks by dynamic A* algorithm straightly.

3.2.1 Updating the backtracking list: Backtracking points are the points which have more than one unvisited neighbouring neurons, that is to say, the potential backtracking point can be a starting point for next coverage path. Backtracking list is a list of backtracking points which are updated as the robot moves. Once the robot moves, the states of the eight neighbours update. An unvisited point will be marked as backtracking point and added into the backtracking list. As shown in Algorithm 1 (see Fig. 2).

```

Input: CurrentPoint(x,y)
Output: the new backtracking list
1 for NeighbourPoints around CurrentPoint do
2   for NeighbourPoints around CurrentPoint do
3     Check State;
4     if State(NeighbourPoints)= Unvisited then
5       UpdateBacktrackinglist(AddCurrentPoint);
6     if all State(NeighbourPoints)= Visit or
       Obstacle then
7       UpdateBacktrackinglist(RemoveCurrentPoint);
    
```

Fig. 2 Algorithm 1: Updating backtracking List. Neighbour points are the 8 points around the current point (x, y). State (position) denotes the state of a neuron at the position. Activity (position) returns the neuron activity of this position. Select (backtracking list) selects the best point for backtracking list. Move (current, target) is a point-to-point planner.

4. APPLICATIONS IN ROBOTS

4.1. NNBased Robotic Manipulator Control.

Generally speaking, the control methods for robot manipulators can be roughly divided into two groups, model-free control and model based control. For the model-free control approaches like proportional-integral-derivative (PID) control, satisfactory control performance may not be guaranteed. In contrast, the model based control approaches exhibit better control behavior but heavily depend on the validity of the robot model. In practice, however, a perfect robotic dynamic model is always not available due to the complex mechanisms and uncertainties. Additionally, the payload may be varied according to different tasks, which makes the accurate dynamics model hard to be obtained in advance. To solve such problems, the NN approximation-based control methods have been used extensively in applications of robot manipulator control. A basic structure of the adaptive neural network control for robot manipulator. Consider a dynamic model of a robot manipulator given as follows:

$$M(q)\ddot{q} + C(\dot{q}, q)\dot{q} + G(q) = \tau,$$

4.2. NN Based Human-Robot Interaction Control.

Recently, there is a predominant tendency to employ the robots in the human-surrounded environment, such as household services or industrial applications, where humans and robots may interact with each other directly. Therefore, interaction control has become a promising research field and has been widely studied. In a learning method was developed such that the dynamics of a robot arm could follow a target impedance model with only knowledge of the robotic structure (see Figure 3).

The NN was further employed in robot control in interaction with an environment where impedance control was achieved with the completely unknown robotic dynamics. In a learning method was developed such that the robot was able to adjust the impedance parameters when it interacted with unknown environments. In order to learn optimal impedance parameters in the robot manipulator control, an adaptive dynamic programming (ADP) method was employed when the robot interacted with unknown time-varying environments, where NNs were used for both critic and actor networks. The ADP was also employed for coordination of multi robots in which possible disagreement between different manipulators



was handled and dynamics of both robots and the manipulated object were not required to be known.

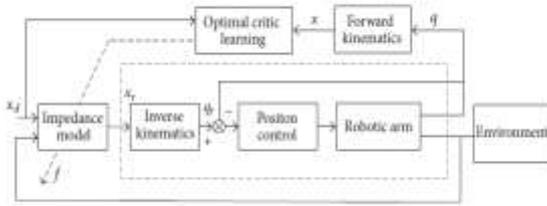


Figure 3: Block of the learning impedance control.

4.3. NN Based Robot Cognitive Control. According to the predictive processing theory the human brain is always actively anticipating the incoming sensorimotor information. This process exists because the living beings exhibit latencies due to neural processing delays and a limited bandwidth in their sensorimotor processing. To compensate for such a delay, in human brain, neural feedback signals (including lateral and top-down connections) modulate the neural activities via inhibitory or excitatory connections by influencing the neuronal population coding of the bottom-up sensory driven signals in the perception-action system. Similarly, in robotic systems, it is claimed that such a delay and a limited bandwidth also can be compensated by the predictive functions learnt by recurrent neural models. Such a learning process can be done via only visual processing or in the loop of perception and action.

Based on the hierarchical sensorimotor integration theory, which advocates that action and perception are intertwined by sharing the same representational basis the representation on different levels of sensory perception does not explicitly represent actions; instead, there is an encoding of the possible future percept which is learnt from prior sensorimotor knowledge.

5. CONCLUSION

In summary, great achievements for control design of nonlinear system by means of neural networks have been gained in the last two decades. Despite the impossibility in identifying or listing all the related contributions in this short review, efforts have been made to summarize the recent progress in the area of NN control and its particular applications in the robot learning control, the robot interaction control, and the robot recognition control. In this paper, we have shown that significant progress of NN has been made in control of the nonlinear systems, in solving the optimization problem, in approximating the system dynamics, in dealing with the input nonlinearities, in

human-robot interaction, and in the pattern recognition. Besides, the proposed method is extended to multi-robot systems with a market-based bidding process, and the workloads are deemed more balanced than other multi-robot approaches. For future research, we plan to consider the energy and timing constraints that allow robots to carry limited energy and to complete the coverage task before the deadline.

6. FUTURE WORK

In future research, the proposed framework will be extended to incorporate energy-efficient path planning and time-constrained task scheduling to address practical mission requirements. The integration of decentralized multi-agent learning strategies will be explored to improve scalability and coordination in large robot teams operating under communication constraints. Further validation through real-world experiments and hardware implementation will be conducted to assess robustness under sensor noise and environmental uncertainties. The inclusion of edge-computing support and cloud-based coordination mechanisms may enhance real-time decision-making capabilities. Additionally, future studies will investigate heterogeneous multi-robot collaboration and safety-aware control strategies to strengthen reliability in dynamic and complex environments.

REFERENCES

1. Nasser, M. Pattern Recognition and Machine Learning. *J. Electron. Imaging* **2007**, 16, 4.
2. Yu, D.; Li, D. Deep learning and its applications to signal and information processing [exploratory dsp]. *IEEE Signal Process. Mag.* **2010**, 28, 145–154. [[CrossRef](#)]
3. Hinton, G.; Deng, L.; Yu, D.; Dahl, G.E.; Mohamed, A.R.; Jaitly, N.; Kingsbury, B. Deep neural Networks for acoustic modeling in speech recognition. *IEEE Signal Process. Mag.* **2012**, 29, 82–97. [[CrossRef](#)]
4. Graves, A.; Abdel-rahman, M.; Geo_rey, H. Speech recognition with deep recurrent neural networks. In *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, Canada, 26–31 May 2013; pp. 6645–6649.
5. Kumar, A. Ask me anything: Dynamic memory networks for natural language processing. In *Proceedings of the International Conference on*



Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 1378–1387.

6. Machine Learning and Natural Language Processing. Available online: <http://l2r.cs.uiuc.edu/~danr/Teaching/CS546-13/Papers/marquez-LNLP00.pdf> (accessed on 11 July 2010).
7. Manning, C. The Stanford CoreNLP natural language processing toolkit. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, MD, USA, 22–27 June 2014; pp. 55–60.
8. Collobert, R.; Jason, W. A unified architecture for natural language processing: Deep neural networks with multitask learning. In Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, 5–9 July 2008; pp. 160–167.
9. Kononenko, I. Machine learning for medical diagnosis: History, state of the art and perspective. *Artif. Intell. Med.* **2001**, *23*, 89–109. [[CrossRef](#)]
10. Shvets, A.A.; Rakhlin, A.; Kalinin, A.A.; Iglovikov, V.I. Automatic instrument segmentation in robot-assisted surgery using deep learning. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; pp. 624–628.