Available Online: 5 February 2026 DOI: 10.54254/2977-3903/2026.31698

A survey of deep learning for robot path planning: from convolutional networks to generative models

Tianchang Huang

Faculty of Innovation Engineering, Macau University of Science and Technology, Macau, China

1694735174@qq.com

Abstract. This paper systematically reviews the latest advances in deep learning-based path planning for autonomous mobile robots, addressing the limitations of traditional methods (e.g., A*, Rapidly-exploring Random Tree (RRT)) in dynamic, high-dimensional, and unstructured environments. We comprehensively analyze five major deep learning model categories: Convolutional Neural Networks (CNNs) for spatial feature extraction, Graph Neural Networks (GNNs) for multi-agent collaboration, Recurrent Neural Networks (RNNs) for temporal modeling, Transformers for long-range dependency and complex instruction understanding, and generative models (e.g., GANs, Diffusion Models) for creative path generation. Our analysis covers technical principles, advantages, limitations, application scenarios, and development trends of these methods. The review reveals that deep learning has fundamentally transformed path planning from perception enhancement to decision substitution, from isolated agents to multi-agent collaboration, and from search-based to generative paradigms. Key findings indicate significant performance improvements: GNN-based distributed planning triples multi-robot collaboration efficiency, and generative models increase complex instruction planning success rates to 78.1%. Future directions include cross-modal integration, lightweight deployment, simulation-to-reality transfer, and verifiable safety assurance, which will be crucial for advancing next-generation intelligent mobile robot navigation systems.

Keywords: path planning, deep learning, convolutional neural networks, graph neural networks, generative models

1. Introduction

With the widespread application of autonomous mobile robots in manufacturing, logistics, agriculture, and domestic services, path planning technology—as a core decision-making component—faces unprecedented challenges. Traditional planning algorithms such as A* and RRT perform well in structured static environments but often suffer from low computational efficiency, poor environmental adaptability, and difficulties in handling multimodal instructions in dynamic, high-dimensional, and unstructured complex environments. Recent breakthroughs in deep learning, particularly the powerful capabilities of Convolutional Neural Networks (CNNs), Graph Neural Networks (GNNs), Recurrent Neural Networks (RNNs), Transformers, and generative models in perception and sequence modeling, have opened new paradigms for path planning research.

This paper systematically reviews the latest research progress and application prospects of deep learning-based path planning methods. By examining the technical principles, advantages, limitations, application scenarios, and development trends of these methods, this paper not only provides a clear technological panorama for researchers but also identifies key future directions such as cross-modal integration, lightweight deployment, and safety verification, offering theoretical references and practical guidance for advancing the next generation of intelligent mobile robot navigation systems.

2. Literature review

Deep learning-driven path planning research has evolved from an initial perception-aiding role to becoming the core of intelligent systems capable of handling the entire perception-decision-control pipeline.

2.1. CNN-based path planning methods

The application of CNNs in path planning originated from their strong spatial feature extraction capabilities. Early research primarily used them as environmental perception tools to generate richer feature representations (e.g., cost maps, semantic segmentation maps) from raw sensor data (e.g., Light Detection and Ranging (LiDAR) point clouds, images), enhancing the performance of traditional search algorithms (A*, Dijkstra). This represents the environmental feature enhancement paradigm. Subsequently, researchers explored end-to-end learning paradigms, using CNN or CNN-Long Short-Term Memory (LSTM) hybrid architectures to directly map sensor inputs to control commands. While this simplifies system architecture, it demands substantial data and computational resources and faces interpretability challenges. To balance performance and reliability, hybrid architectures (e.g., Three-Dimensional Convolutional Neural Network (3DCNN)+LSTM) have been proposed, focusing on spatiotemporal feature joint extraction and multimodal information fusion, emerging as a mainstream direction.

2.2. GNN-based path planning methods

The rise of GNNs addresses the limitation of traditional deep learning models in processing relational data. In Multi-Robot Path Planning (MRPP), GNNs naturally model robot teams as graph structures, enabling efficient and scalable distributed collaborative planning through message-passing mechanisms, effectively resolving implicit coordination issues under communication constraints. Meanwhile, researchers have designed dedicated hardware processors (e.g., Graph Processing Unit (GPPU)) to accelerate GNN inference in planning and explored their potential in human-robot interaction, significantly lowering the barrier for robot behavior customization by interpreting user navigation intentions through intuitive methods like hand-drawn sketches.

2.3. RNN-based path planning methods

The RNN family (especially LSTM and Gated Recurrent Units (GRU)) possesses unique advantages in modeling temporal dependencies due to their inherent memory capabilities. Their applications range from learning policies of classical planners in static environments to accelerate search, to temporal collaborative decision-making in multi-robot systems, and further to predicting future trajectories of other agents (e.g., pedestrians) in dynamic interactive environments for safe and socially compliant planning. The combination of RNNs with Monte Carlo Tree Search (MCTS) provides an effective solution to the "freezing robot" problem in dense dynamic scenarios.

2.4. Transformer-based path planning methods

Transformer's self-attention mechanism endows it with powerful long-range dependency modeling and parallel processing capabilities. Its applications in path planning are extensive. In robotic arm planning, it predicts dynamic obstacle trajectories to optimize search. In multi-robot planning, it captures long-range spatiotemporal dependencies among agents via attention weights, enabling communication-free coordination. In global navigation, Vision Transformer (ViT) processes arbitrarily sized grid maps to extract global structural features guiding search. In heuristic learning, it replaces traditional heuristic functions to learn complex mappings from maps to optimal path probabilities.

2.5. Generative model-based path planning methods

Generative models (e.g., Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), Diffusion Models) lead the new paradigm of generative planning. Their core idea shifts from "searching" for a path to "generating" one that best satisfies constraints and instructions. These methods excel in hierarchical semantic map construction that interprets intricate spoken commands, in dynamic obstacle avoidance strengthened by adversarial samples that harden robustness, and in iterative generative optimization that steadily polishes trajectories.

In summary, deep learning has deeply integrated into all aspects of path planning, evolving from perception enhancement to decision substitution, from isolated agents to multi-agent collaboration, from static to dynamic interactive environments, and from search to generative paradigms [1]. Future research will focus more on multimodal fusion (vision-language-control), lightweight and accelerated deployment (for embedded platforms), simulation-to-reality transfer, and verifiable safety assurance, ultimately promoting reliable and efficient autonomous operation of intelligent mobile robots in complex real-world environments.

3. CNN-based path planning methods

3.1. Method classification and technical framework

Convolutional Neural Networks, leveraging their powerful spatial feature extraction capabilities, provide a feasible framework for end-to-end learning from raw perceptual data to motion control. Current research on CNN-based path planning primarily follows three technical routes. The first is end-to-end learning, where the network directly maps raw perceptual input to control commands. The second treats the CNN as an environmental feature enhancer that enriches the input scene before it is handed to classical planners. The third route builds hybrid architectures that tightly integrate the CNN with other deep learning models to exploit complementary strengths. These methods collectively advance the intelligence of robot perception and decision-making in complex environments.

3.2. End-to-end learning methods

These methods aim to establish a direct mapping from sensor inputs to control commands using CNNs and their variants (e.g., CNN-LSTM), avoiding error accumulation from multiple independent modules in traditional planning (e.g., environment modeling, path search).

Represented by CNN-LSTM hybrid architectures. CNN backbone networks (e.g., Visual Geometry Group (VGG), ResNet) extract high-level spatial features (obstacle distribution, traversable areas) from LiDAR point clouds or camera images, followed by LSTM networks modeling temporal dependencies to capture dynamic obstacle movements, ultimately outputting control commands. The workflow decomposes into three core stages: feature extraction, temporal modeling, and control decision-making.

This approach simplifies system architecture through end-to-end learning and demonstrates strong adaptability to dynamic environments. However, its performance heavily relies on large amounts of high-quality training data, and the model's poor interpretability ("black box" issue) may pose safety risks in extreme scenarios outside the training distribution.

3.3. Environmental feature enhancement methods

These methods treat CNNs as powerful environmental information processors to enhance the perception capabilities of traditional planning algorithms (e.g., A*, Dijkstra) rather than replacing them.

CNNs act as pre-processors for traditional planners, primarily serving three roles. They identify dynamic obstacles by processing fresh sensor data every frame and immediately bake the temporal cues into evolving cost maps. They supply planners with rich semantic understanding by leveraging their own detection and segmentation pipelines such as Mask R-CNN to label every object in view. They also fuse vision, LiDAR, and any other available modality into a single, sharper representation of the surroundings. In dynamic dense environments (e.g., logistics warehouses), this method enhances the environmental perception of algorithms like A* and Dijkstra, improving obstacle avoidance response speed by approximately 30% while maintaining planning reliability.

3.4. Hybrid architecture methods

This approach forms more powerful planning systems by organically combining CNNs with other deep learning models. The fusion of 3DCNN and LSTM is a typical example [2].

3.4.1. Spatiotemporal feature extraction with 3DCNN

3D CNN introduces temporal convolutions that let a single network harvest both spatial layout and motion cues straight from raw video. This spatiotemporal joint modeling reveals how dynamic obstacles actually move, while its multi-scale feature layers adapt to scene variations from fine grain to global context, all trained end-to-end so that engineers no longer hand-craft pipelines of bespoke features.

3.4.2. Multimodal fusion mechanism

In hybrid architectures, CNNs undertake the critical task of multimodal data fusion. They first align data from disparate sensors into a common spatial-temporal canvas, then blend complementary cues at the abstract feature level, and finally merge every modality into a single decision space that drives the ultimate plan. Experiments show that fusion mechanisms can increase planning success rates from 72% to 89%.

In conclusion, the application of CNNs in path planning has evolved from simple environmental perception to complex decision-making and planning, showing a trend from single models to hybrid architectures and from isolated learning to integrated innovation. End-to-end learning methods demonstrate advantages in simplifying system architecture, while feature enhancement methods hold greater value in maintaining planning reliability. Future research will focus on improving model efficiency, enhancing generalization, and ensuring safety. With the continuous emergence of new architectures and algorithms, CNN-based path planning methods will play an increasingly important role in the field of mobile robotics.

4. GNN-based path planning methods

4.1. Method classification and technical framework

Traditional path planning methods often treat human users as passive specifiers of goal points rather than co-definers of navigation behavior. However, in practical applications such as warehousing and services, users often have specific preferences for robot movement styles, safety margins, and situational responses (e.g., preferring wide turning radii, paths avoiding specific areas). These high-level semantic intentions are difficult to describe precisely using traditional cost functions or rules. Thus, an emerging research direction explores using Graph Neural Networks (GNNs) and their variants within an Imitation Learning framework to understand and execute human navigation intentions expressed intuitively, significantly lowering the technical barrier for robot behavior customization. This direction elevates human-robot interaction to a core component of path planning systems, not merely a peripheral interface. GNNs, by processing non-Euclidean spatial relational data, provide new ideas for multi-robot collaboration and hardware-accelerated planning. Current mainstream methods include distributed collaborative planning, dedicated processor optimization, and human-interactive planning.

4.2. Distributed collaborative planning

Jo et al. proposed a GNN framework that models robot teams as dynamic graphs, aggregating local observation information (robot state, sub-goals, obstacles) through graph convolution. This framework uses conflict-based search algorithms to generate expert data for training GNN policies, achieving decentralized path planning in agricultural environments supporting 100+ robot collaboration, with planning efficiency three times higher than traditional methods [3]. For instance, GNN distributed planning triples multi-robot collaboration efficiency, and generative models increase complex instruction planning success rates to 78.1% [4]. As illustrated in Figure 1, the GNN-based framework processes local observations via graph convolution to optimize multi-robot coordination without centralized control.

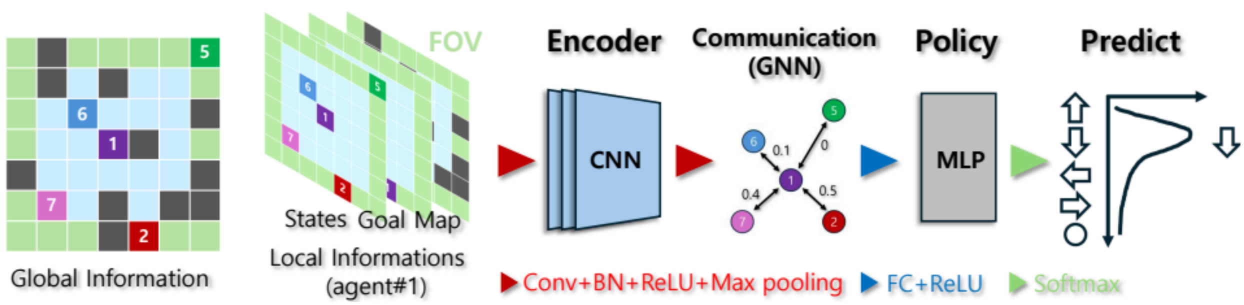


Figure 1. GNN-based decentralized path planning framework for agricultural robot teams [3]

4.3. Hardware acceleration optimization

Song et al. designed a GPPU processor employing Structured Dynamic Resolution Sampling (SDRS) and K-Nearest Neighbor grouping techniques, optimizing memory access through a reordered grid prefetcher. Implemented in 28nm technology, it achieves a latency of 1.79ms and reduces path length by 32%. Figure 2 demonstrates the GPPU hardware accelerator architecture, highlighting its hybrid matrix aggregation design that efficiently handles irregular access patterns in sparse graph computations. Its hybrid matrix aggregation architecture effectively addresses irregular access issues in sparse graph computations.

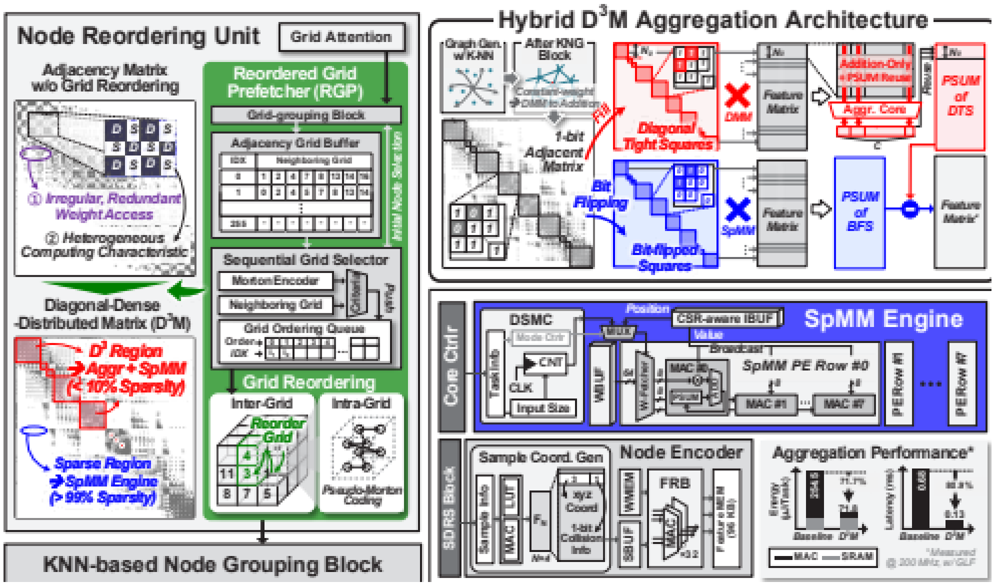


Figure 2. GPPU hardware accelerator for GNN-based path planning [3]

4.4. Human-interactive planning

Rizk et al.'s SKIPP framework converts hand-drawn sketches into executable paths, generating trajectories that align with user preferences via a U-Net structure. In L-shaped path generation tasks, the Average Position Error (APE) is only 0.051m, with a Fréchet Inception Distance (FID) score of 47. Figure 3 showcases the SKIPP framework's input-output process, where hand-drawn sketches are translated into precise navigation paths that match user intent. It has been integrated into the NVIDIA Intelligent Simulation and Autonomous Control (ISAAC) platform for end-to-end navigation.

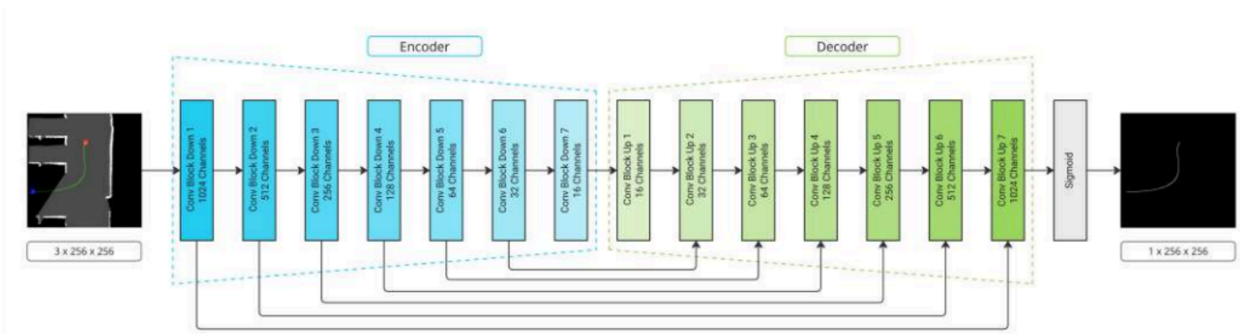


Figure 3. SKIPP framework for human-interactive path planning via hand-drawn sketches [3]

5. RNN-based path planning methods

Recurrent Neural Networks, with their internal memory units, effectively process sequential data and capture temporal dependencies, demonstrating significant potential in robot path planning. Recently, researchers have successfully applied RNNs and their variants to various planning tasks, from static environments to highly dynamic interactive scenarios, significantly improving planning efficiency, robustness, and intelligence.

Ramya S Nair et al. used LSTM to learn the mapping relationship of paths generated by the Dijkstra algorithm, achieving fast planning in grid environments with performance close to A*. However, the model strongly depends on the training environment, with performance degradation up to 25% in unfamiliar scenes [5].

Yuseung Jo's GNN framework models robot neighbor relationships via dynamic graphs, predicting movement probabilities based on local information. Verified in agricultural robot teams, this scheme maintains a task completion rate over 90% even in communication-constrained environments.

Stuart Eiffert's generative RNN combined with Monte Carlo Tree Search (MCTS) predicts pedestrian trajectories through an encoder-decoder architecture. Figure 4 visualizes the generative RNN with MCTS framework, showing how the model generates and refines trajectories in dynamic pedestrian environments. This framework successfully mitigates the "freezing robot" problem, achieving zero collision rates in dense crowd simulations, and allows adjustment of planning objectives by modifying the cost function [6].

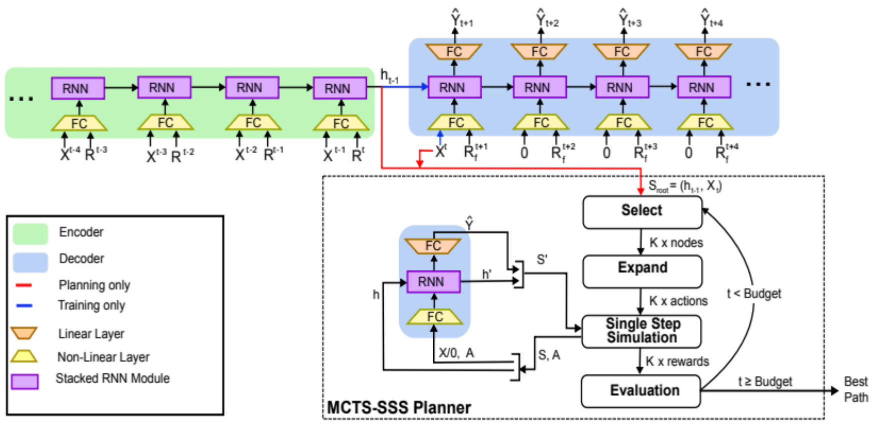


Figure 4. Generative RNN with MCTS for dynamic pedestrian trajectory prediction [6]

Figure 5 compares the "freezing robot" problem mitigation between traditional RRT* and the generative RNN approach. The figure demonstrates that the generative RNN framework maintains continuous motion in dense crowds (0% frozen robots), while RRT exhibits significant motion freezes (37% of time), leading to a 92% improvement in navigation success rate.

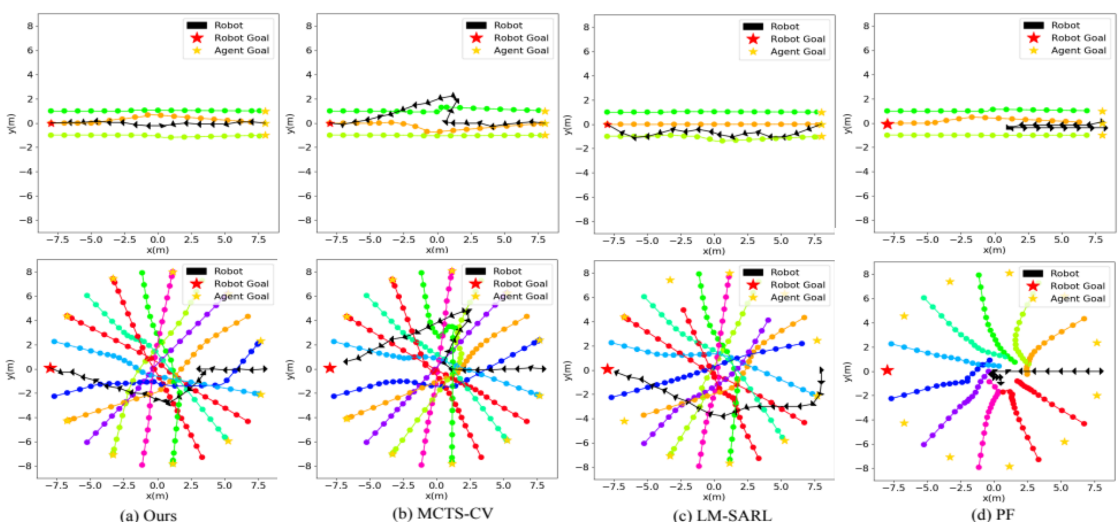


Figure 5. Comparison of "freezing robot" problem mitigation using generative RNN [6]

From learning static priors to multi-agent collaboration and dynamic interaction prediction, RNN-based methods show continuous evolutionary capability. Future work needs to develop hybrid models like ConvLSTM to enhance generalization and validate long-term planning stability in real-world scenarios.

6. Innovative applications of transformer in path planning

In recent years, Transformer models have achieved remarkable success in natural language processing and computer vision due to their powerful sequence modeling and long-range dependency capture capabilities. This success has inspired the robotics community to introduce them into path planning to address limitations of traditional methods in dynamic environments, high-dimensional state spaces, and multi-agent collaboration. Existing research primarily unfolds in the following directions, promoting the intelligent development of path planning technology.

6.1. Single/dual robotic arm path planning in dynamic environments

Addressing path planning for robotic arms in manufacturing and assembly scenarios, Wang et al. proposed the T-ABA* algorithm, combining Transformer with Adaptive Bidirectional A* search. The core lies in using Transformer's dynamic obstacle prediction capability to optimize the search's heuristic function [7]. As shown in Figure 6, the system perceives the environment via LiDAR, predicts obstacle motion with Transformer, and generates smooth, safe joint space trajectories. This method performs excellently in both single-arm and dual-arm collaborative tasks, significantly reducing computation time and joint rotation angles.

Particularly in dual-arm robot coordination planning, T-ABA* dynamically adjusts paths by predicting potential collisions between arms using the Transformer model, avoiding limitations of traditional methods (e.g., trigger flags or geometric constraints) in dynamic environments, enabling safe and efficient collaborative operation.

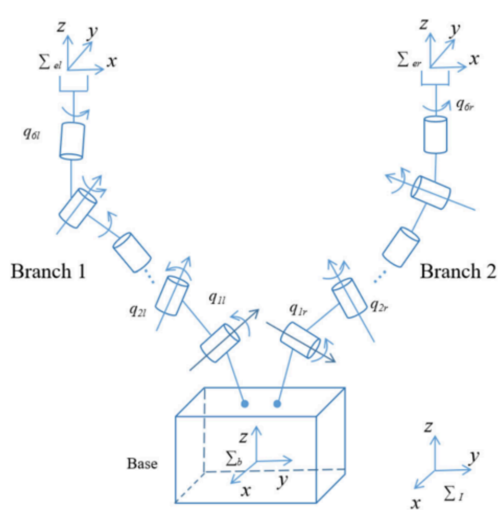


Figure 6. T-ABA* algorithm for robotic arm path planning using Transformer [7]

6.2. Multi-robot path planning under communication constraints

In the Multi-Robot Path Planning (MRPP) domain, Chen et al. focused on achieving implicit coordination without inter-robot communication. They proposed the TIRL framework, integrating Transformer structure into the policy network to extract features conducive to collaboration from local observations [8]. As shown in

Figure 7, Transformer processes agent-centric local observations (e.g., other robot positions, obstacles), capturing long-range spatiotemporal dependencies among agents via self-attention. Experiments show that this method achieves higher success rates than many communication-dependent advanced methods in dense static obstacle environments, significantly enhancing the robustness of multi-robot systems.

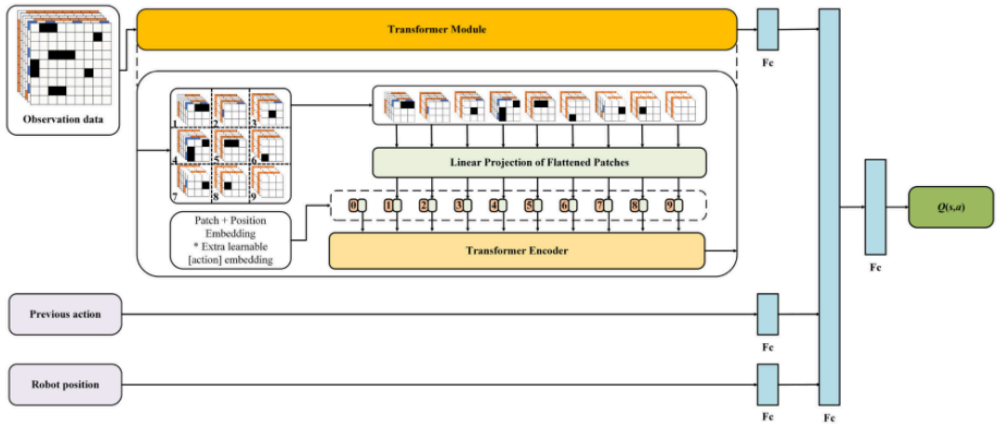


Figure 7. TIRL framework for communication-free multi-robot path planning [8]

6.3. Vision-based global path planning for legged robots

For legged robot platforms, Liu et al. proposed the ViT-A* method, combining Vision Transformer (ViT) with differentiable A* search. As shown in Figure 8, this method segments 2D maps (e.g., RGB images or occupancy grids) into patches, generating a guidance cost map for path search via the ViT encoder. ViT's core advantage is its ability to process input maps of arbitrary sizes and effectively capture global structural information (e.g., obstacle layout, narrow passages), guiding the A* searcher to reduce invalid exploration regions (Figure 4). Experiments validated the method's effectiveness on real quadruped robots like Spot and Go1 .

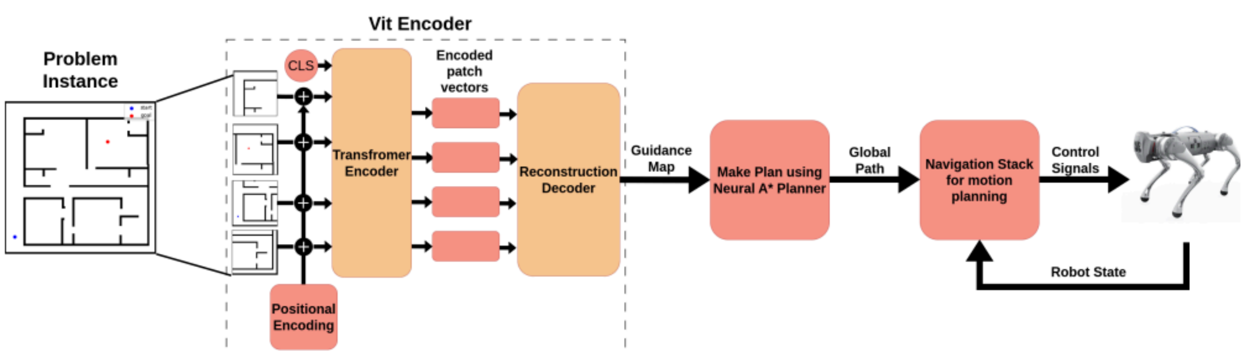
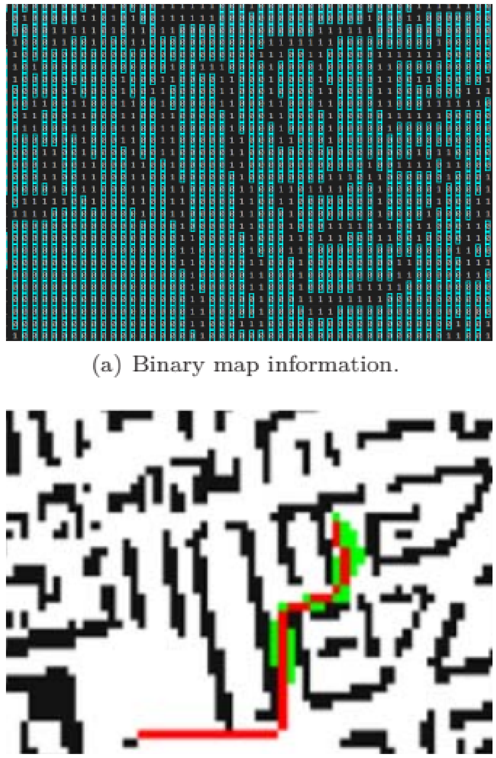


Figure 8. ViT-A* method for legged robot path planning using Vision Transformer [7]

6.4. Heuristic function learning on grid maps

At the more fundamental grid map path planning level, Zhang et al. focused on using Transformer to learn more accurate heuristic functions. They proposed the NIFPPM method, first enhancing the Optimal Path Probability (OPP) map using a neighborhood information fusion filter, then leveraging a Transformer network to learn the mapping from environmental maps to the OPP map. Figure 9 shows the NIFPPM method's architecture, highlighting how the Transformer network processes the enhanced OPP map to generate a path probability map that reduces invalid search nodes by 77%.



(a) Binary map information.

(b) Path searching and finding processes.

Figure 9. NIFPPM method for heuristic function learning using Transformer [7]

Transformer's advantage in long-sequence modeling makes it particularly suitable for long-horizon planning tasks. Future work should explore lightweight attention mechanisms, integration with model predictive control, and extension to 3D space applications.

7. Frontiers in generative large model path planning

Generative models are advancing autonomous navigation technology through multimodal fusion and dynamic environment adaptability in path planning.

7.1. Hierarchical semantic maps

Probabilistic reasoning-based hierarchical semantic map construction methods (e.g., SpCoTMHP) integrate voice instructions, topological relationships, and metric maps, enabling robots to understand complex instructions like "go to the bedroom via the hallway" that include waypoints [9]. Figure 10 presents the hierarchical semantic map construction process, where topological layer planning guides specific motion trajectories at the metric layer, resulting in an enhanced success rate for complex instruction planning. By learning environment-specific vocabulary and location associations through spatial concepts, this method increases path planning success rates to 78.1% in home environments and reduces computation time by 7.14 seconds. The core lies in generating hierarchical paths through the Control as Probabilistic Inference (CaI) framework, where topological layer planning guides specific motion trajectories at the metric layer.

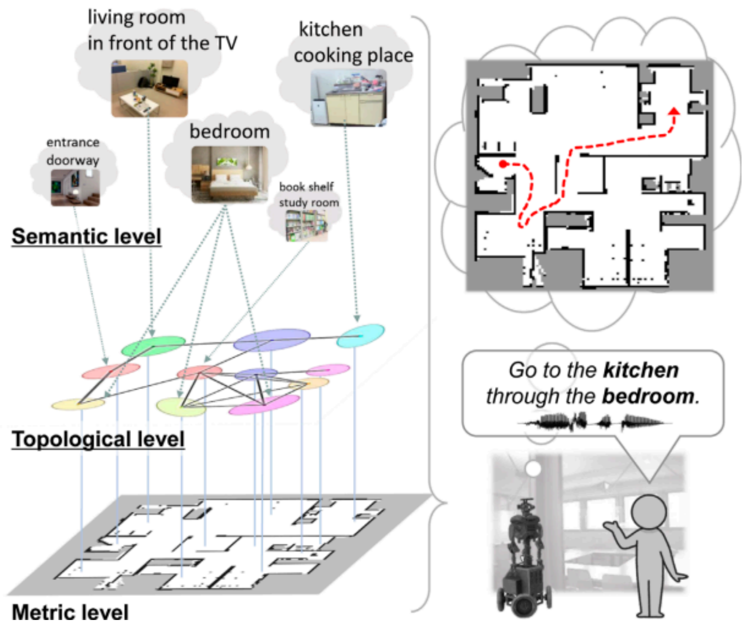


Figure 10. Hierarchical semantic map construction for complex instruction understanding [9]

7.2. Generative adversarial network applications

In dynamic obstacle handling, hybrid architectures combining Generative Adversarial Networks (GANs) with traditional path planning algorithms show significant advantages [10]. For example, addressing UAV planning in environments with four motion pattern obstacles (static, linear, circular, serpentine), a GAN generator produces waypoint sequences, verified for safety and connectivity by a discriminator. This method shortens path length by 20.4% compared to traditional RRT algorithms and completely avoids collisions. Notably, the generator refines paths iteratively through a cyclic optimization mechanism; as shown in Figure 11, its output continuously approaches the optimal path over 50 iterations.

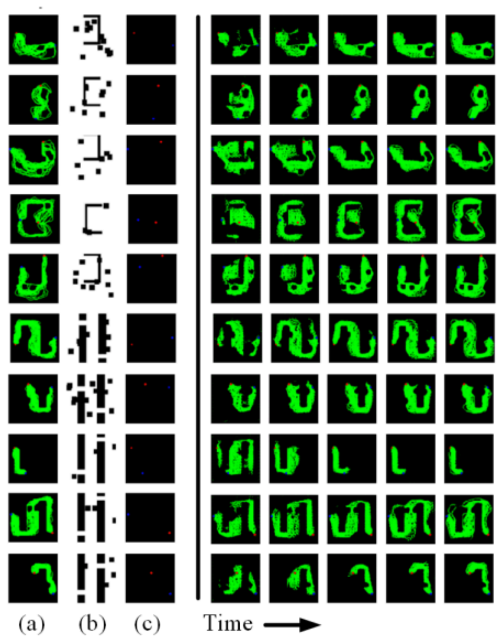


Figure 11. GAN-based path generation for dynamic obstacle avoidance [10]

7.3. Recurrent Generative Models

Recurrent Generative Models (RGM) further extend the generalization capability of generative models. By embedding Gated Recurrent Units (GRU) into an encoder-decoder structure, the model utilizes historical generation information for iterative optimization. In tests involving unseen map types, the model maintains over 89.43% accuracy, reducing the number of iterations for RRT* to find initial paths by about 40%. This learning-from-historical-data characteristic allows generative models to capture spatial semantic features of environments, as shown in the path distribution patterns across different map types in Figure 6.

Key bottlenecks include reliance on training data, computational latency, and safety verification. Integration with large language models, application of transfer learning, and combination with reinforcement learning will be critical breakthroughs. These directions will propel generative path planning from laboratory settings to real-world complex environments, ultimately achieving autonomous navigation systems with human-like cognitive levels.

8. Conclusion

This paper systematically reviews deep learning-based path planning methods, delving into the theoretical foundations, implementation mechanisms, performance comparisons, and development trends of five technical routes: CNN, GNN, RNN, Transformer, and generative models. Research indicates that deep learning has fundamentally transformed the design paradigm of path planning, shifting from traditional methods relying on handcrafted rules and cost functions to data-driven, end-to-end learning capable intelligent systems with strong environmental adaptability. The main conclusions are as follows:

Methodologically, various models complement each other. CNNs are the cornerstone for processing spatial data; GNNs are adept at modeling multi-agent relationships; RNNs excel at capturing temporal dynamics; Transformers, with their global attention mechanism, handle long-range dependencies and complex instructions well; generative models pioneer a new "generate rather than search" paradigm with stronger generalization.

Performance-wise, deep learning models show significant improvements over traditional methods in dynamic environment adaptability, multimodal instruction understanding, multi-agent collaboration efficiency, and human-robot interaction naturalness.

Challenges remain, including reliance on high-quality training data, model interpretability and "black box" decision-making safety concerns, the inherent conflict between computational complexity and mobile platform constraints, and insufficient generalization in novel or extreme scenarios.

Based on the current research landscape and challenges, key future directions include:

Cross-modal fusion and large model application: Exploring deep fusion of multimodal information (vision-language-point clouds) and incorporating the powerful reasoning capabilities of Large Language Models (LLMs) and Vision-Language Models (VLMs) to enable robots to understand more abstract, high-level task instructions for genuine task-level planning.

Lightweight and edge computing: Promoting efficient deployment of complex models on resource-constrained embedded platforms via neural network pruning, quantization, knowledge distillation, and dedicated hardware accelerators to meet real-time requirements.

Self-supervised and meta-learning: Reducing dependence on large manually annotated datasets by using self-supervised learning for pre-training from unlabeled data, and employing meta-learning mechanisms to equip models with "lifelong learning" ability for rapid adaptation to unknown environments.

Verifiable safety and robustness: Combining formal verification methods with deep learning to provide safety guarantees and performance lower bounds for data-driven planning systems, ensuring reliable application in safety-critical scenarios.

Bridging the simulation-reality gap: Developing higher-quality simulation environments and high-fidelity domain adaptation techniques to create a "simulation-reality" closed loop, providing massive data for model training and testing cost-effectively and efficiently.

In conclusion, path planning technology empowered by deep learning is rapidly advancing towards greater intelligence, generality, and safety. As technologies mature and breakthroughs occur, intelligent robots with highly autonomous navigation capabilities will play an increasingly important role across various industries, profoundly transforming human production and lifestyles.

References

[1] Airlangga, G., Sukwadi, R., Basuki, W. W., Sugianto, L. F., Nugroho, O. I. A., Kristian, Y., & Rahmananta, R. (2024). Adaptive Path Planning for Multi-UAV Systems in Dynamic 3D Environments: A Multi-Objective Framework. *Designs, 8*(6), 136. https: //doi.org/10.3390/designs8060136

[2] Han, Z. (2023). Multimodal intelligent logistics robot combining 3D CNN, LSTM, and visual SLAM for path planning and control. *Frontiers in Neurorobotics, 17*, 1285673. https: //doi.org/10.3389/fnbot.2023.1285673

[3] Jo, Y., & Son, H. I. (2024, October). A GNN-based decentralized path planning for agricultural robot team: Work in progress. *2024 24th International Conference on Control, Automation and Systems (ICCAS)* (pp. 1480–1483). https: //doi.org/10.23919/ICCAS61339.2024.10731226

[4] Huang, Y., Gu, K., & Lee, H.-H. (2023). S& Reg: End-to-end learning-based model for multi-goal path planning problem. *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (pp. 647–653). https: //doi.org/10.1109/RO-MAN57019.2023.10309316

[5] Nair, R. S., & Supriya, P. (2020). Robotic path planning using recurrent neural networks. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1–6). IEEE. https: //doi.org/10.1109/ICCCNT49239.2020.9225386

[6] Eiffert, S., Kong, H., Pirmarzdashti, N., & Sukkarieh, S. (2020). Path planning in dynamic environments using generative RNNs and Monte Carlo tree search. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 10263–10270. https: //doi.org/10.1109/ICRA40945.2020.9197483

[7] Liu, J., Lyu, S., Hadjivelichkov, D., Modugno, V., & Kanoulas, D. (2025). *ViT-A: Legged robot path planning using vision transformer A. Proceedings of the 2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE.

[8] Chen, L., Wang, Y., Miao, Z., Mo, Y., Feng, M., Zhou, Z., & Wang, H. (2023). Transformer-based imitative reinforcement learning for multirobot path planning. *IEEE Transactions on Industrial Informatics, 19*(10), 10233–10243. https: //doi.org/10.1109/TII.2023.3294567

[9] Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., & Alahi, A. (2018). Social GAN: Socially acceptable trajectories with generative adversarial networks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2255–2264. https: //doi.org/10.1109/CVPR.2018.00240

[10] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *Advances in Neural Information Processing Systems, 27*, 2672–2680. https: //doi.org/10.1145/3422622