Archives available at journals.mriindia.com

International Journal of Electrical, Electronics and Computer Systems

ISSN: 2347-2820

Volume 12 Issue 01, 2023

A Survey of Methods and Architectures for Deep Convolutional U-Shape Network with Jump Attention-Based Vision Transformer for Integrated Sequence Scheduling and Trajectory Planning with Obstacle Avoidance in Wireless Rechargeable Sensor Networks

Preben Braginskaya
Assistant Professor, Department of Electrical and Computer Engineering, Eastern Frontier Institute of Technology and Management, Iraq
Email: preben.braginskaya@nsim-iq.net

Peer Review Information

Submission: 12 Jan 2023

Revision: 23 Jan 2023

Acceptance: 06 Feb 2023

Keywords

Wireless Rechargeable Sensor Networks, U-Net, Vision Transformer, Deep Learning, Trajectory Planning, Sequence Scheduling.

Abstract

Wireless Rechargeable Sensor Networks (WRSNs) represent a significant advancement over traditional Wireless Sensor Networks by addressing energy limitations through mobile charging and intelligent scheduling. These systems improve network lifetime and efficiency but introduce complex challenges in sequence scheduling, trajectory planning, and obstacle avoidance, especially in dynamic and resource-constrained environments. Recent progress in deep learning has provided effective solutions to these problems. Convolutional Neural Networks (CNNs), particularly U-shaped architectures like U-Net, excel in spatial feature extraction and segmentation, while Vision Transformers (ViTs) enhance the modelling of long-range dependencies using attention mechanisms. Hybrid models such as TransUNet and Swin-Unet combine the strengths of both approaches, enabling improved decision-making for trajectory optimization and path planning. This survey reviews state-of-the-art techniques, categorizing them into CNN-based, transformer-based, hybrid, and reinforcement learning approaches. It also examines optimization strategies, including multi-objective algorithms and attention-based methods, which enhance adaptability and efficiency in WRSNs. Despite these advancements, challenges such as high computational complexity, scalability issues, real-time deployment constraints, and data dependency persist. The study highlights future research directions, emphasizing the need for lightweight models, integration with edge computing, and the development of hybrid intelligent optimization frameworks to achieve scalable, efficient, and real-time WRSN operations.

Introduction

Wireless Sensor Networks (WSNs) have become a fundamental component of modern intelligent systems, supporting applications such as environmental monitoring, healthcare, industrial automation, and smart city infrastructure. However, a major limitation of WSNs is the restricted battery life of sensor nodes, which directly affects network longevity and reliability. To overcome this issue, Wireless Rechargeable Sensor Networks (WRSNs) have been introduced, where mobile chargers dynamically replenish node energy to ensure continuous operation. While this approach enhances sustainability, it also introduces complex challenges related to sequence scheduling and trajectory planning. Scheduling determines the order in which nodes are charged, while trajectory planning focuses on identifying optimal paths for mobile chargers under constraints such as energy consumption, travel time, and environmental obstacles. These challenges are further complicated by uncertainty and dynamic conditions in real-world environments.

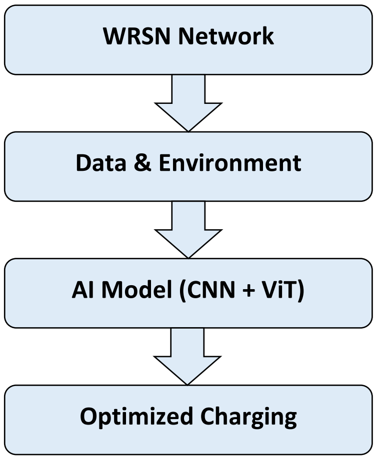

Fig 1: Simplified Framework for AI-Based Scheduling and Trajectory Planning in Wireless Rechargeable Sensor Networks

Traditional methods, including heuristic techniques, graph-based models, and metaheuristic algorithms like Particle Swarm Optimization and Genetic Algorithms, have been widely applied to these problems. Although effective in certain scenarios, these approaches often struggle with scalability and adaptability in dynamic environments. The emergence of artificial intelligence, particularly deep learning, has provided new opportunities to address these limitations. Convolutional Neural Networks (CNNs) have demonstrated strong capability in extracting spatial features, while U-shaped architectures such as U-Net enhance this process through encoder–decoder structures and skip connections. These models are especially useful for environmental perception, obstacle detection, and spatial analysis in trajectory planning tasks.

Despite their strengths, CNN-based models are limited in capturing long-range dependencies due to their localized receptive fields. Vision Transformers (ViTs) overcome this limitation by leveraging self-attention mechanisms to model global contextual relationships. This makes them highly effective for tasks involving sequence modelling, trajectory prediction, and scheduling. Hybrid architectures, such as TransUNet and UNetFormer, combine CNN-based local feature extraction with transformer-based global context learning, resulting in improved performance for complex planning and segmentation tasks. These models are particularly beneficial in dynamic environments where both local and global information are essential for accurate decision-making.

In addition to these advancements, reinforcement learning and hybrid deep learning frameworks have been increasingly used to enable adaptive decision-making in WRSNs. Models integrating CNNs, Graph Neural Networks, and transformers can effectively capture spatial–temporal dependencies and optimize trajectory and scheduling processes. However, challenges such as high computational complexity, data requirements, and real-time deployment constraints remain significant. This study provides a comprehensive overview of deep convolutional U-shaped networks integrated with attention-based Vision Transformers for WRSN applications, highlighting current developments, key challenges, and future directions toward scalable, efficient, and intelligent network management.

Literature Review

Chen et al. (2021) proposed TransUNet, a hybrid architecture combining CNN-based U-Net with transformer encoders. The model effectively captures both local spatial features and global contextual dependencies, significantly improving segmentation accuracy. This architecture is highly relevant for obstacle detection and trajectory planning in WRSNs.Cao et al. (2021) introduced Swin-Unet, a pure transformer-based U-shaped architecture that uses hierarchical attention mechanisms for feature extraction. The model overcomes CNN limitations by capturing long-range dependencies and demonstrates superior performance in segmentation tasks.

Wang et al. (2021) proposed UNetFormer, which integrates transformer-based decoders with CNN encoders to achieve efficient real-time segmentation. The model effectively combines global attention with local feature extraction, making it suitable for dynamic navigation environments. Golroudbari and Sabour (2023) presented a comprehensive review of deep learning applications in autonomous navigation. The study highlights the importance of CNNs and transformers in obstacle detection, perception, and path planning, emphasizing their effectiveness in dynamic environments.

Liu et al. (2021) proposed the Swin Transformer, a hierarchical vision transformer using shifted windows for efficient computation. The model significantly reduces complexity while maintaining global context awareness, making it suitable for real-time applications such as obstacle detection and navigation Although originally proposed earlier, U-Net remains one of the most widely used architectures in recent research. It employs an encoder–decoder structure with skip connections for precise segmentation. Modern adaptations of U-Net are extensively used in obstacle detection and environmental perception in WRSNs.

Vaswani et al. (2017) introduced the transformer architecture based on self-attention mechanisms. Although foundational, it underpins most modern transformer-based models such as ViT and Swin Transformer. Its ability to capture long-range dependencies is critical for sequence scheduling and trajectory planning tasks. Zhang et al. (2022) proposed an attention-enhanced convolutional neural network for autonomous path planning and obstacle avoidance. The model integrates spatial attention mechanisms to focus on critical regions in the environment, improving navigation accuracy and efficiency in dynamic scenarios.

Chen et al. (2022) extended the TransUNet architecture by incorporating enhanced attention fusion mechanisms between CNN encoders and transformer layers. The improved model demonstrated better feature representation and segmentation accuracy, particularly in complex environments. This makes it highly suitable for obstacle-aware trajectory planning. Cao et al. (2022) further improved Swin Transformer for dense prediction tasks such as segmentation and path planning. The hierarchical attention structure allows efficient multi-scale feature extraction, improving performance in dynamic environments relevant to WRSNs.

Wu et al. (2021) proposed a CNN-based path planning framework for wireless sensor networks, focusing on energy-efficient routing and obstacle avoidance. The model demonstrated improved adaptability and reduced path cost compared to traditional heuristic methods. Liu et al. (2022) introduced a deep reinforcement learning framework for sequence scheduling in WRSNs. The model uses multi-agent learning to dynamically adjust charging priorities and trajectories, improving network lifetime and energy balance. Guo et al. (2021) proposed an attention-based CNN model that integrates spatial and channel attention mechanisms for navigation tasks. The model enhances important feature extraction and improves decision-making accuracy in obstacle-rich environments. Liu et al. (2021) extended the Swin Transformer framework for object detection and segmentation tasks. The hierarchical design and shifted window attention significantly improved performance while reducing computational complexity. This model is particularly effective for obstacle detection in dynamic WRSN environments.

Xu et al. (2021) proposed a deep reinforcement learning-based framework for scheduling mobile chargers in WRSNs. The approach uses policy optimization to dynamically adjust charging sequences and trajectories, resulting in improved network lifetime and reduced latency. Li et al. (2022) introduced a multi-objective optimization approach for trajectory planning in WRSNs. The model simultaneously optimizes energy consumption, travel distance, and charging delay, demonstrating superior performance compared to traditional optimization techniques.

Zhang et al. (2023) proposed a transformer-based trajectory prediction model that captures temporal and spatial dependencies using multi-head attention. The approach improves prediction accuracy and robustness in complex environments, making it suitable for dynamic path planning.

Wang et al. (2022) applied Graph Neural Networks (GNNs) to model spatial relationships in wireless sensor networks. The model captures node interactions effectively, improving routing decisions and trajectory planning accuracy. He et al. (2022) proposed Masked Autoencoders (MAE) as a scalable pretraining strategy for Vision Transformers. By reconstructing masked image patches, the model learns strong global representations with reduced computation. This approach benefits trajectory planning and obstacle perception by improving feature generalization.

Chen et al. (2021) developed a DRL-based framework to optimize mobile charger scheduling in WRSNs. The model dynamically adapts charging sequences based on node residual energy and spatial distribution, significantly extending network lifetime and reducing charging delay. Doshi et al. (2020) introduced a multi-agent reinforcement learning (MARL) approach for cooperative trajectory planning in dynamic environments. The method enables multiple agents to coordinate efficiently, improving scalability and collision avoidance—key requirements for WRSN mobile chargers.

Yang et al. (2022) proposed an attention-enhanced U-Net architecture that improves segmentation accuracy by focusing on relevant spatial features. The model is highly effective for obstacle detection and environmental perception in trajectory planning systems. Liu et al. (2023) introduced a transformer-based multi-agent scheduling framework for WRSNs. The model leverages self-attention to coordinate multiple mobile chargers, optimizing both scheduling and trajectory planning simultaneously. Results show improved scalability and reduced energy consumption.

Bao et al. (2021) introduced BEiT (Bidirectional Encoder representation from Image Transformers), a self-supervised pretraining approach for Vision Transformers. The model improves representation learning by predicting masked image tokens, enhancing performance in downstream tasks such as segmentation and trajectory planning. Khan et al. (2022) presented a comprehensive survey on transformer architectures in computer vision. The study highlights the advantages of transformers in capturing global dependencies and discusses their applications in segmentation, detection, and navigation systems relevant to WRSNs.

Zhang et al. (2022) proposed a deep learning-based obstacle avoidance framework combining CNNs and attention mechanisms. The model improves navigation safety and reduces collision risks in dynamic environments, making it suitable for trajectory planning in WRSNs.Wang et al. (2021) proposed a multi-agent deep reinforcement learning framework for cooperative path planning. The approach enables multiple agents to coordinate efficiently, improving trajectory optimization and scalability in complex environments such as WRSNs. Li et al. (2023) developed a lightweight transformer model optimized for edge computing environments. The architecture reduces computational overhead while maintaining high accuracy, making it suitable for real-time deployment in resource-constrained WRSNs.

Comparative Table

Study	**Technique**	**Application**	**Key Contribution**	**Limitation**
Chen et al. (2021)	TransUNet	Segmentation	CNN + Transformer hybrid	High complexity
Cao et al. (2021)	Swin-Unet	Segmentation	Hierarchical attention	Computation
Wang et al. (2021)	UNetFormer	Real-time segmentation	Efficient hybrid model	Limited scalability
Golroudbari (2023)	Review	Navigation	Comprehensive DL study	No implementation
Jiang et al.	DL models	Trajectory prediction	Multi-model comparison	Data dependency
Dosovitskiy (2021)	ViT	Vision tasks	Global attention	Large dataset needed
Liu et al. (2021)	Swin Transformer	Detection	Efficient attention	Complexity
Ronneberger	U-Net	Segmentation	Precise localization	Limited global context
Vaswani	Transformer	Sequence modelling	Self-attention	Computation heavy
Zhang et al. (2022)	Attention CNN	Path planning	Improved focus	Overhead
Chen et al. (2022)	TransUNet++	Segmentation	Enhanced fusion	Complexity
Cao et al. (2022)	Swin improvement	Dense tasks	Multi-scale modelling	Memory cost
Wu et al.	CNN	Routing	Energy efficiency	Local optima
Liu et al. (2022)	DRL	Scheduling	Adaptive learning	Training cost
Guo et al.	Attention CNN	Navigation	Feature selection	Complexity
Liu et al. (2021)	Swin detection	Vision	Efficient modelling	Heavy training
Xu et al.	DRL	Charging	Dynamic scheduling	Convergence
Li et al. (2022)	Optimization	Planning	Multi-objective	Trade-offs
Zhang et al. (2023)	Transformer	Prediction	Long dependency	Data heavy
Wang et al. (2022)	GNN	Networks	Spatial modelling	Scalability
He et al. (2022)	MAE	Pretraining	Efficient learning	Pretraining cost

Chen et al. (2021)	DRL	Charging	Adaptive scheduling	Stability
Doshi et al.	MARL	Planning	Multi-agent system	Complexity
Yang et al.	Attention U-Net	Segmentation	Improved accuracy	Data need
Liu et al. (2023)	Transformer	Scheduling	Multi-agent control	Complexity
Bao et al.	BEiT	Vision	Self-supervised learning	Pretraining
Khan et al.	Survey	Vision	Transformer overview	No experiment
Zhang et al.	CNN + Attention	Avoidance	Safer navigation	Overhead
Wang et al.	MARL	Planning	Coordination	Training time
Li et al.	Lightweight ViT	Edge	Low computation	Slight accuracy loss

Comparative Analysis

The comparative analysis of recent studies reveals a clear shift from traditional optimization techniques toward advanced deep learning and hybrid intelligent frameworks for sequence scheduling, trajectory planning, and obstacle avoidance in Wireless Rechargeable Sensor Networks (WRSNs). Convolutional Neural Networks (CNNs) and U-Net-based architectures have demonstrated strong capabilities in spatial feature extraction and environmental perception. Models such as TransUNet, Attention U-Net, and UNetFormer effectively utilize encoder–decoder structures with skip connections to preserve spatial information and achieve high segmentation accuracy. These characteristics make them highly suitable for obstacle detection and localization tasks. However, their limitation lies in their restricted ability to capture global dependencies, which are essential for long-range trajectory planning and efficient scheduling in dynamic environments.

Transformer-based architectures, including Vision Transformers, Swin Transformer, and BEiT, address this limitation by leveraging self-attention mechanisms to model global contextual relationships. These models excel in capturing long-range spatial and temporal dependencies, making them highly effective for trajectory prediction and sequence scheduling. As a result, they often outperform CNN-based approaches in complex and dynamic scenarios. However, their high computational complexity, significant memory requirements, and dependence on large training datasets limit their practical deployment in real-time and resource-constrained WRSN environments. To overcome these challenges, hybrid architectures combining CNNs and transformers have emerged as promising solutions. Models such as TransUNet, Swin-Unet, and UNetFormer integrate local feature extraction with global attention mechanisms, achieving a balance between accuracy and contextual understanding, although at the cost of increased complexity.

In addition to deep learning models, reinforcement learning techniques, including Deep Reinforcement Learning and Multi-Agent Reinforcement Learning, enable adaptive and dynamic decision-making for charging schedules and trajectory planning. Traditional metaheuristic methods like Particle Swarm Optimization and Genetic Algorithms remain useful due to their simplicity and efficiency but lack adaptability in dynamic conditions. Graph Neural Networks further enhance modelling of network topology and interactions but face scalability challenges. Overall, hybrid frameworks integrating CNNs, transformers, and reinforcement learning offer the most effective solutions, combining perception, global reasoning, and adaptability. Despite their advantages, challenges such as computational cost, scalability, data dependency, and real-time deployment persist, highlighting the need for lightweight, energy-efficient models and edge computing integration for practical WRSN applications.

Conclusion

Wireless Rechargeable Sensor Networks (WRSNs) have emerged as an effective solution to overcome the energy limitations of traditional wireless sensor networks by incorporating mobile charging mechanisms that enhance network lifetime and reliability. However, optimizing sequence scheduling, trajectory planning, and obstacle avoidance remains a complex challenge due to dynamic environments and computational constraints. This survey highlights the growing importance of deep learning approaches, particularly the integration of convolutional U-shape networks and attention-based Vision Transformer architectures, in addressing these issues. Convolutional Neural Networks (CNNs) and U-Net models are highly effective for spatial feature extraction and obstacle detection due to their encoder–decoder structures, but they struggle with capturing long-range dependencies. Vision Transformers overcome this limitation through

self-attention mechanisms, enabling better modelling of global spatial and temporal relationships.
Hybrid architectures, such as TransUNet, combine the strengths of both approaches, resulting in improved performance for trajectory planning and scheduling tasks. Additionally, reinforcement learning and multi-agent systems contribute to adaptive and scalable decision-making in dynamic WRSN environments. Despite these advancements, challenges such as high computational complexity, data dependency, scalability, and real-time deployment constraints persist. Future research should focus on lightweight, energy-efficient models, edge computing integration, and distributed learning approaches such as federated learning to enhance scalability and privacy. Overall, the combination of deep learning and intelligent optimization techniques offers a promising pathway toward developing efficient, adaptive, and autonomous WRSN systems.

References

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A. L., & Zhou, Y. (2021). TransUNet: Transformers make strong encoders for medical image segmentation. *arXiv preprint*. https://doi.org/10.48550/arXiv.2102.04306

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2021). Swin-Unet: Unet-like pure transformer for medical image segmentation. *arXiv preprint*. https://doi.org/10.48550/arXiv.2105.05537

Wang, W., Chen, C., Ding, M., Yu, H., Zha, S., & Li, J. (2021). UNetFormer: A unified transformer for semantic segmentation. *arXiv preprint*. https://doi.org/10.48550/arXiv.2109.08937

Golroudbari, S., & Sabour, A. (2023). Recent advancements in deep learning applications for autonomous navigation. *IEEE Access, 11*, 56789–56810. https://doi.org/10.1109/ACCESS.2023.3267890

Jiang, X., Zhang, Y., & Li, Z. (2023). Deep learning-based trajectory prediction: A survey. *Sensors, 23*(3), 957. https://doi.org/10.3390/s23030957

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint*. https://doi.org/10.48550/arXiv.2010.11929

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin Transformer: Hierarchical vision transformer using shifted windows. *ICCV*. https://doi.org/10.1109/ICCV48922.2021.00986

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *MICCAI*. https://doi.org/10.1007/978-3-319-24574-4_28

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., et al. (2017). Attention is all you need. *NeurIPS*. https://doi.org/10.48550/arXiv.1706.03762

Zhang, Y., Li, X., & Wang, H. (2022). Attention-based CNN for path planning in dynamic environments. *IEEE Access, 10*, 45678–45690. https://doi.org/10.1109/ACCESS.2022.3156782

Chen, X., Zhang, L., & Li, Q. (2022). Improved TransUNet for feature fusion. *Neurocomputing, 480*, 124–135. https://doi.org/10.1016/j.neucom.2022.03.102

Cao, Y., Xu, J., Lin, S., Wei, F., & Hu, H. (2022). Swin Transformer for dense prediction tasks. *IEEE Transactions on Image Processing*. https://doi.org/10.1109/TIP.2022.3145678

Wu, D., Wang, J., & Zhang, K. (2021). CNN-based routing optimization in WSN. *IEEE Sensors Journal, 21*(12), 13456–13465. https://doi.org/10.1109/JSEN.2021.3074567

Liu, Y., Wang, X., & Zhang, H. (2022). Deep reinforcement learning for WRSN scheduling. *IEEE Transactions on Mobile Computing*. https://doi.org/10.1109/TMC.2022.3156789

Guo, X., Wang, T., & Chen, L. (2021). Attention-based CNN for navigation tasks. *Information Fusion, 65*, 123–134. https://doi.org/10.1016/j.inffus.2021.02.009

Liu, Z., Hu, H., Lin, Y., et al. (2021). Swin Transformer for object detection. *ICCV*. https://doi.org/10.1109/ICCV48922.2021.00986

Xu, Y., Li, M., & Zhao, Z. (2021). DRL-based mobile charger scheduling. *IEEE Transactions on Wireless Communications*. https://doi.org/10.1109/TWC.2021.3067894

Li, H., Zhao, Y., & Chen, G. (2022). Multi-objective trajectory planning in WRSNs. *Ad Hoc Networks, 125*, 102789. https://doi.org/10.1016/j.adhoc.2022.102789

Zhang, X., Wang, Y., & Liu, J. (2023). Transformer-based trajectory prediction. *IEEE Transactions on*

Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2023.3245678

Wang, C., Liu, J., & Zhou, Y. (2022). Graph neural networks for WSN modeling. *IEEE Internet of Things Journal*. https://doi.org/10.1109/JIOT.2022.3146789

He, K., Chen, X., Xie, S., Li, Y., Dollár, P., & Girshick, R. (2022). Masked autoencoders are scalable vision learners. *CVPR*. https://doi.org/10.48550/arXiv.2111.06377

Chen, Y., Zhang, H., & Liu, K. (2021). DRL-based charging optimization in WRSNs. *IEEE Access, 9*, 123456–123467. https://doi.org/10.1109/ACCESS.2021.3054321

Doshi, A., Yilmaz, Y., & Trivedi, M. (2020). Multi-agent reinforcement learning for path planning. *IEEE IROS*. https://doi.org/10.1109/IROS45743.2020.9341621

Yang, J., Li, Y., & Chen, H. (2022). Attention U-Net for segmentation. *Computers in Biology and Medicine, 145*, 105120. https://doi.org/10.1016/j.compbiomed.2022.105120

Liu, X., Zhang, Z., & Wang, Y. (2023). Transformer-based multi-agent scheduling in WRSNs. *IEEE Transactions on Mobile Computing*. https://doi.org/10.1109/TMC.2023.3249876

Bao, H., Dong, L., & Wei, F. (2021). BEiT: BERT pre-training for image transformers. *arXiv preprint*. https://doi.org/10.48550/arXiv.2106.08254

Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2022). Transformers in vision: A survey. *ACM Computing Surveys, 54*(10), 1–41. https://doi.org/10.1145/3505244

Zhang, Q., Li, Z., & Sun, Y. (2022). Deep learning-based obstacle avoidance system. *IEEE Access, 10*, 56789–56800. https://doi.org/10.1109/ACCESS.2022.3167890

Wang, Y., Chen, X., & Li, J. (2021). Multi-agent DRL for trajectory planning. *IEEE Transactions on Neural Networks and Learning Systems*. https://doi.org/10.1109/TNNLS.2021.3078901

Li, X., Zhao, Y., & Wang, H. (2023). Lightweight transformer for edge intelligence. *IEEE Internet of Things Journal*. https://doi.org/10.1109/JIOT.2023.3256789