Computer Engineering and Intelligent Systems www.iiste.org
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.15, No.1, 2024

A NOVEL FRAMEWORK TO EFFICIENT PATH PLANNING THROUGH REAL-TIME COST MAP GENERATION USING NEURAL NETWORKS FOR SEARCH AND RESCUE MISSIONS.

Keith Machina
Department of Computer Science, Kent State University
kmachina@kent.edu

Jong-Hoon Kim
Department of Computer Science, Kent State University
jkim72@kent.edu

Awal Ahmed
Department of Computer Science, Kent State University
afime@kent.edu

Abstract

Search and rescue (SAR) missions are critical operations that demand swift and efficient execution to save lives in the aftermath of dis- asters. This paper introduces a novel framework for optimizing path planning in robotic SAR missions through the generation of real-time cost maps using neural networks. Our approach integrates static topo- logical data with dynamic mission findings to create an amalgamated cost map that prioritizes urgent and accessible regions. We propose a modified U-Net architecture, specifically adapted for SAR applications, which enables adaptive cost prediction and enhances learning capabilities in complex, evolving environments. Extensive simulations demonstrate significant improvements in survivor location efficiency compared to traditional baseline approaches. The framework's ability to continuously update based on real-time data ensures robust adapt- ability to the dynamic nature of SAR missions. By bridging the gap between theoretical models and practical implementation, our method has the potential to revolutionize crisis response strategies, offering a more agile and effective approach to robotic search and rescue operations.
DOI: 10.7176/CEIS/15-1-10
Publication date: October 30th 2024

1 Introduction

In the wake of natural disasters or man-made calamities, the efficiency of search and rescue (SAR) operations can mean the difference between life and death for survivors trapped in hazardous environments. As the com- plexity and scale of these disasters continue to grow, there is an increasing need for advanced technological solutions to augment human efforts in SAR missions. The integration of robotics and artificial intelligence into these operations has shown great promise, offering the potential to navigate dan- gerous terrains, process vast amounts of data, and make rapid decisions in time-critical situations (Laporte, 1992).

However, the deployment of robotic systems in SAR scenarios presents unique challenges that go beyond traditional path planning problems. The dynamic nature of disaster environments, coupled with the urgent need to locate survivors quickly, creates a complex optimization problem that de- fies conventional solutions. Factors such as changing terrain conditions, the discovery of new information during the mission, and the need to prioritize certain areas over others all contribute to the intricacy of the task.

This research addresses these challenges by proposing a novel framework for efficient path planning in robotic SAR missions. At its core, our approach treats the problem as a variant of the traveling salesman problem (TSP), but with critical modifications to account for the unique aspects of SAR operations. We introduce a hybrid technique that synergizes both algorithmic and non-algorithmic approaches, leveraging the power of neural networks to generate real-time cost maps for efficient decision-making.

The key contributions of this paper are threefold:

- We present a framework for generating real-time amalgamated cost maps that seamlessly integrate static topological data with dynamic mission findings. This novel approach allows for a more nuanced and adaptive path planning strategy that can respond to changing condi- tions and priorities during the mission.
- We introduce a modified U-Net architecture tailored specifically for SAR applications. This adaptation enables adaptive cost prediction and enhances the learning capabilities of the system in complex, evolv- ing environments.
- We provide a comprehensive evaluation of our proposed framework against established baseline approaches in a series of simulated SAR scenarios. These simulations demonstrate the superior performance of our method in terms of survivor location efficiency and overall mission effectiveness.

By addressing the limitations of existing methods, our approach aims to significantly enhance the efficiency and effectiveness of robotic SAR op- erations. The potential impact of this research extends beyond theoretical advancements, offering practical solutions that could lead to faster survivor recovery and improved outcomes in real-world crisis scenarios.

In the following sections, we delve into the related work that forms the foundation of our research, provide a detailed explanation of our methodol- ogy, present the results of our simulations, and discuss the implications of our findings. We conclude by outlining future research directions and the po- tential applications of our framework in broader contexts of disaster response and crisis management.

2 Related Work

The optimization of path planning for search and rescue (SAR) operations draws from a rich body of research across multiple disciplines, including robotics, artificial intelligence, and operations research. At its core, the prob- lem shares similarities with the classical Traveling Salesman Problem (TSP), which has been extensively studied in computer science and mathematics. However, the dynamic and time-critical nature of SAR missions introduces additional complexities that require novel approaches.

2.1 Algorithmic Approaches to Path Planning

Traditional solutions to TSP-like problems can be broadly categorized into algorithmic and non-algorithmic approaches. Algorithmic techniques, char- acterized by their systematic and deterministic nature, have seen significant advancements in recent years (Held & Karp, 1971). These methods follow well-defined procedures to find optimal or near-optimal solutions, but often struggle with the computational complexity inherent in NP-Hard problems like the TSP, especially when applied to the dynamic environments typical of SAR scenarios.

Several notable algorithmic techniques have pushed the boundaries of TSP solutions:

- The Lin-Kernighan Heuristic has proven to be one of the most effective methods for generating near-optimal solutions to the TSP (Rosenkrantz, Stearns, & Lewis, 1977). This algorithm works by making a series of edge exchanges to iteratively improve the tour. Its success in finding high-quality solutions has made it a benchmark against which other TSP algorithms are often compared.
- Ant Colony Optimization, inspired by the foraging behavior of ants, has shown promising results in solving TSP instances (Gutin, Punnen, Barvinok, Gimadi, & Serdyukov, 2001). This meta-heuristic approach uses artificial ants to construct solutions, depositing pheromones on edges to guide future iterations towards better solutions. The adapt- ability of this method makes it particularly interesting for dynamic environments.
- Genetic Algorithms, drawing inspiration from the principles of natural selection and evolution, have been successfully applied to TSP and its variants (Applegate, Bixby, Chv´atal, & Cook, 2006). These algorithms evolve a population of solutions over time, using techniques such as crossover and mutation to generate improved solutions. The ability of genetic algorithms to maintain a diverse set of solutions makes them robust in changing environments.
- The Branch and Bound algorithm, an exact method for solving op- timization problems, has seen improvements through various bound- ing techniques and parallelization (Cormen, Leiserson, Rivest, & Stein, 2009; Johnson & McGeoch, 1997). While traditionally limited by its computational requirements, recent

advancements have expanded its applicability to larger problem instances.

Despite these advancements, algorithmic approaches often struggle with the real-time adaptability required in SAR missions. The dynamic nature of disaster environments, where new information constantly emerges and priorities shift rapidly, poses significant challenges to these methods.

2.2 Non-Algorithmic Approaches and Machine Learn- ing

In response to the limitations of purely algorithmic methods, researchers have increasingly turned to non-algorithmic techniques, particularly those rooted in machine learning and artificial intelligence. These approaches of- fer greater flexibility and adaptability, crucial qualities in the unpredictable environments characteristic of SAR operations.

Markov Decision Processes (MDPs) and Reinforcement Learning (RL) have emerged as promising frameworks for decision-making in complex, un- certain environments (Niroui, Zhang, Kashino, & Nejat, 2019). MDPs pro- vide a mathematical framework for modeling decision-making in situations where outcomes are partly random and partly under the control of a decision- maker. This aligns well with the uncertain nature of SAR missions, where the outcomes of actions are not always predictable.

Reinforcement Learning, particularly advanced techniques like Deep Q- Networks (DQN), offers a powerful approach to learning optimal policies through interaction with the environment. The ability of RL algorithms to adapt to changing conditions and learn from experience makes them par- ticularly attractive for SAR applications. However, these methods also face several challenges in practical implementations:

- Computational Complexity: As the state space of the problem grows, MDPs can suffer from the "curse of dimensionality," making them com- putationally intractable for large-scale problems (Kober, Bagnell, & Pe- ters, 2013). This is particularly problematic in SAR scenarios, where the state space can be vast and complex.

- Sample Efficiency: Many RL algorithms require a large number of inter- actions with the environment to learn effective policies (Cully, Clune, Tarapore, & Mouret, 2015). In time-critical SAR missions, where each decision carries significant weight, the luxury of extensive trial-and- error learning is often not available.

- Exploration-Exploitation Dilemma: Balancing the need to explore new areas (to potentially find better solutions) with exploiting known good strategies is a fundamental challenge in RL (Zhu et al., 2016). In SAR contexts, where time is of the essence, this balance becomes even more critical and difficult to manage.

•

2.3 Hybrid Approaches and Neural Networks

The limitations of both purely algorithmic and purely learning-based meth- ods have led researchers to explore hybrid approaches that combine the strengths of multiple techniques. Neural networks, with their ability to learn complex patterns and make rapid predictions, have become a cornerstone of many such hybrid solutions. The U-Net architecture, originally developed for biomedical image seg- mentation (Ronneberger, Fischer, & Brox, 2015), has shown remarkable adaptability to various domains requiring spatial understanding and local- global context integration. Its encoder-decoder structure, coupled with skip connections, allows for the preservation of both fine-grained details and broader contextual information, making it a promising candidate for gen- erating cost maps in SAR scenarios.

Our research builds upon these foundations, proposing a novel framework that leverages a modified U-Net architecture to generate real-time cost maps for SAR path planning. By combining the pattern recognition capabilities of neural networks with traditional path planning algorithms, we aim to create a system that is both adaptive to changing conditions and computationally efficient enough for real-time deployment.

In the following sections, we detail our methodology, explaining how we've adapted and extended these existing approaches to meet the unique chal- lenges of SAR missions. Our framework represents a step forward in bridging the gap between theoretical advancements and practical, deployable solutions for robotic search and rescue operations.

3 Methodology

Our proposed framework for efficient path planning in search and rescue (SAR) missions comprises several interconnected components, each designed to address specific challenges inherent in these complex operations. The methodology integrates data preparation, cost map generation, neural net- work design, and path planning algorithms to create a comprehensive solution capable of adapting to the dynamic nature of SAR environments.

3.1 Data Preparation and Representation

The foundation of our approach lies in the comprehensive representation of the search area. We model a typical SAR environment as a 64x64 grid, corre- sponding to an area of approximately 0.34 by 0.21 miles. This discretization allows for detailed mapping of various environmental features while main- taining computational feasibility.
Within this grid, we extract and process five key data dimensions, each capturing crucial aspects of the SAR environment:

Table 1: Sample Data Dimensions

Dimension	Description	Range
Road Network	Road occupancy	1 (occupied), 0 (unoccupied), -1 (impassable)
Geographical Features	Physical features	0.1 (most important) to 0.5 (least important)
Slope	Terrain steepness	Assumed 1 for all cells in this study
Survivor Presence	Survivor indicator	1 (present), 0 (absent)
Findings	Evidence relevance	0.1 to 0.5
Report Information	Reported locations	1 (reported), 0 (unreported)

This multi-dimensional representation allows our system to capture both static environmental features (such as road networks and terrain character- istics) and dynamic mission-specific information (like survivor presence and emerging findings). The integration of these diverse data types is crucial for creating a nuanced understanding of the search area, enabling more informed decision-making in path planning.

3.2 Cost Map Generation

Central to our framework is the generation of three types of cost maps, each serving a specific purpose in guiding the SAR operation:

3.2.1 Static Cost Map

The static cost map represents the fixed environmental challenges of travers- ing the search area. It is calculated using a weighted combination of road network data, geographical features, and slope information:

Static Cost = 0.5(Road Network) + 0.4(Geographical Features) + 0.1(Slope)

This formulation prioritizes the accessibility of areas (road network) while also considering the impact of geographical features and terrain steepness. The weights in this equation were determined through empirical testing and domain expertise, balancing the influence of each factor on overall traversability.

3.2.2 Priority Map

The priority map captures the urgency and importance of investigating differ- ent areas within the search grid. It is a dynamic representation that evolves as new information becomes available during the mission. The priority for each cell is determined by the following function:

Priority = { 1 if survivor is present,
0.6(Report Info) + 0.3(Survivor) + 0.1(Findings) if report info present,
0.2(Report Info) + 0.4(Survivor) + 0.4(Findings) otherwise
}

This function ensures that cells with confirmed survivor presence receive the highest priority. In the absence of confirmed survivors, the priority is calculated based on a weighted combination of reported information, potential survivor presence, and relevance of findings. This adaptive prioritization allows the system to focus on the most promising or critical areas as the mission progresses.

3.2.3 Amalgamated Cost Map

The amalgamated cost map is the cornerstone of our path planning strategy. It combines the information from the static cost map and the priority map to create a comprehensive representation of the search area:

$$\text{Amalgamated Cost} = \frac{\text{Priority}}{\text{Static Cost}}$$

This formulation ensures that areas of high priority and low traversal cost are favored in the path planning process. The amalgamated cost map serves as the primary input to our neural network, enabling it to learn and predict optimal paths that balance the urgency of reaching high-priority areas with the practicality of efficient movement through the environment.
Figure 1 illustrates the process of generating these interconnected cost maps.

3.3 Neural Network Design

The heart of our framework is a modified U-Net architecture, specifically adapted for the task of generating real-time amalgamated cost maps in SAR scenarios. The U-Net, originally developed for biomedical image segmentation (?, ?), offers several advantages that make it well-suited for our application:

0.25	0.15
0.435	0.5675

→

0.0123	0.0781
0.0816	0.020

Figure 1: Cost Map Generation Process

- Its encoder-decoder structure allows for the capture of both local de- tails and global context, crucial for understanding the complex spatial relationships in SAR environments.

- Skip connections between the encoder and decoder paths enable the preservation of fine-grained information, ensuring that important de- tails are not lost in the compression process.

- The architecture's ability to work with limited training data aligns well with the constraints often faced in developing SAR systems.

Our modifications to the standard U-Net architecture include:

- Introduction of attention regions: This allows the network to focus more intensely on areas of high importance or complexity within the search grid.

- Removal of overlapping sliding kernels during convolution: This mod- ification reduces computational complexity and aligns better with the goal of identifying priority areas rather than focusing on fine-grained relationships between neighboring cells.

- Adaptation of the network depth and width to balance between model capacity and computational efficiency, ensuring real-time performance in resource-constrained environments.

Figure 2 provides a visual representation of our modified U-Net architecture.

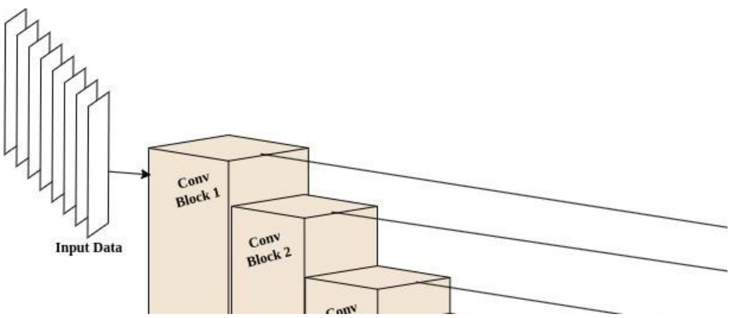

Figure 2: Modified U-Net Architecture for SAR Cost Map Generation

3.4 Path Planning Integration

The final component of our framework involves integrating the generated amalgamated cost map with a path planning algorithm to guide the SAR robot's movements. We employ the A* algorithm, a widely-used and efficient heuristic search algorithm, for this purpose (Laporte, 1992). The A* algorithm is particularly well-suited for our application due to its ability to find the optimal path while considering both the cost of reaching a node and the estimated cost to the goal. In our framework, the amalgamated cost map directly informs the cost function used by A*, ensuring that the algorithm prioritizes paths that balance urgency (high priority areas) with efficiency (low traversal cost).
The integration process works as follows:

1. The neural network generates the amalgamated cost map based on current environmental data and mission findings.

2. The A* algorithm uses this cost map to calculate the optimal path to the next high-priority area.

3. As the robot traverses the chosen path, it continuously collects new data, updating the dynamic aspects of the environment representation.

4. The neural network updates the amalgamated cost map in real-time based on this new information.

5. The process repeats, with A* recalculating the optimal path at regular intervals or when significant new information is acquired.

This continuous feedback loop between environmental sensing, cost map generation, and path planning allows our system to adapt swiftly to changing conditions and new discoveries during the SAR mission.
Figure 3 provides a comprehensive overview of our entire framework, il- lustrating the interconnections between data processing, neural network op- erations, and path planning algorithms.
By combining the pattern recognition capabilities of neural networks with the efficiency of traditional path planning algorithms, our framework aims to provide a robust, adaptive, and computationally feasible solution for real- time SAR operations. The following sections will detail the performance of this framework in simulated scenarios, demonstrating its efficacy compared to baseline approaches.

4 Results and Discussion

To evaluate the efficacy of our proposed framework, we conducted extensive simulations using a variety of scenarios designed to mimic real-world SAR operations. Our primary metric for assessment was the time taken to locate survivors, measured in unit time, where one unit is equivalent to the time required for the robot to traverse one cell in the search grid.

4.1 Baseline Approaches

To provide context for our results, we implemented two baseline approaches:

4.1.1 Naive Approach

This method employed a simple row-by-row search strategy, mimicking an uninformed search pattern. In our simulations, this approach required 4024 unit time to locate all 16 survivors in the training set. This serves as our worst-case scenario, highlighting the inefficiency of uninformed search strate- gies in complex environments.

4.1.2 Static Cost Map

This more advanced baseline used a fixed cost map based solely on unchang- ing environmental features to guide the search. This method showed marked improvement over the naive approach, locating all survivors in 2413 unit time. While this demonstrates the value of incorporating environmental data, it also underscores the limitations of relying on static information in dynamic SAR scenarios.
Figure 4 illustrates the performance of the static cost map approach on the training data.

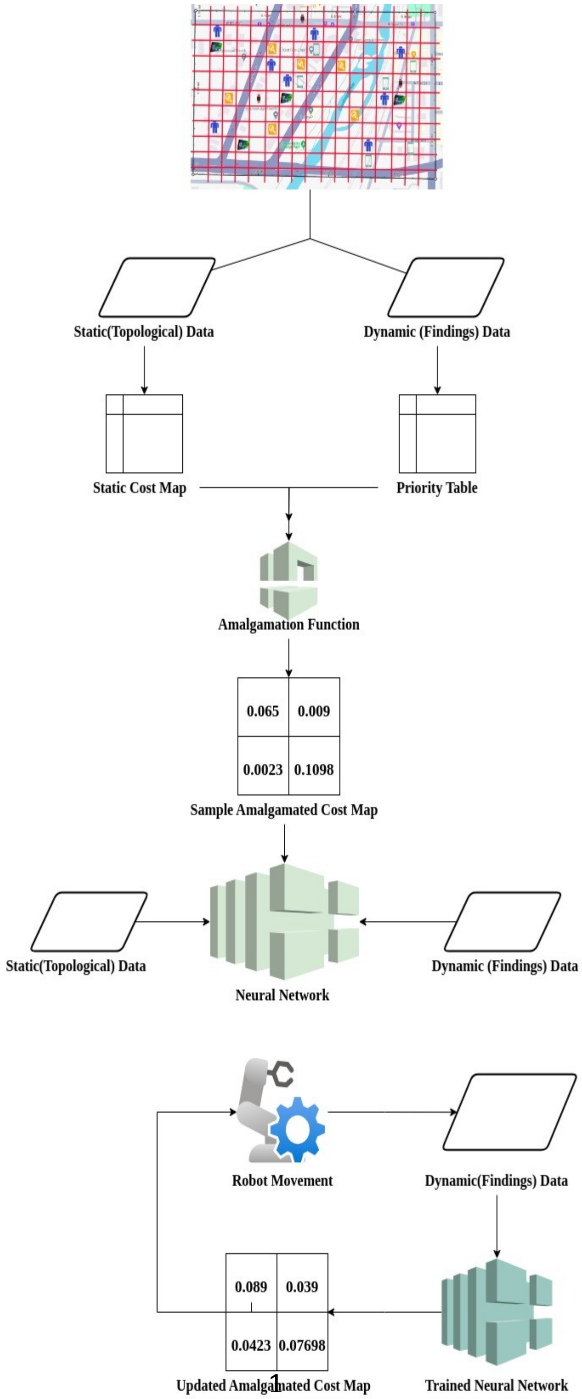

Figure 3: Comprehensive Framework Overview

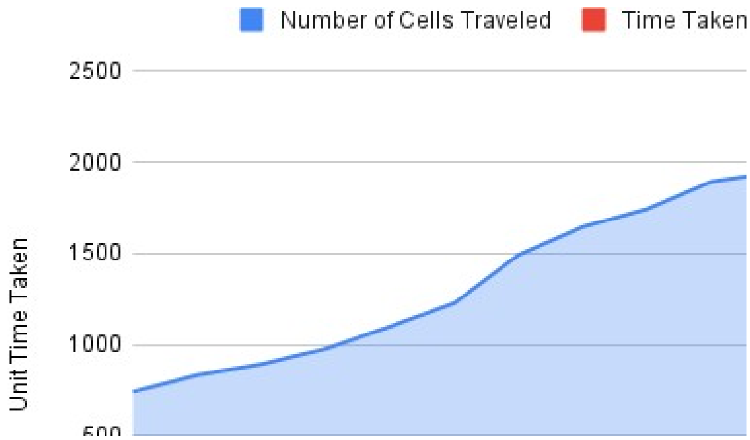

Figure 4: Performance of Static Cost Map Approach on Training Data

4.2 Performance of the Original U-Net

Our initial implementation using the original U-Net architecture (referred to as Model 10) demonstrated significant improvements over the baseline approaches:

- On the training dataset, Model 10 located the first survivor at 495 unit time and all survivors at 1514 unit time.
- When tested on unseen data, it located all survivors at 3170 unit time.
- Stress testing with varying report information release times yielded the following results:
  - Training set averages: First survivor at 489 unit time, all survivors at 1624 unit time.
  - Validation set averages: First survivor at 94 unit time, all sur- vivors at 3140 unit time.

These results showcase the U-Net's ability to effectively integrate both static and dynamic information to guide the search process. The marked improvement in initial survivor detection time is particularly noteworthy, as rapid initial discoveries can be crucial in real-world rescue operations.
Figures 5 and 6 illustrate the performance of the original U-Net on train- ing and test data, respectively.

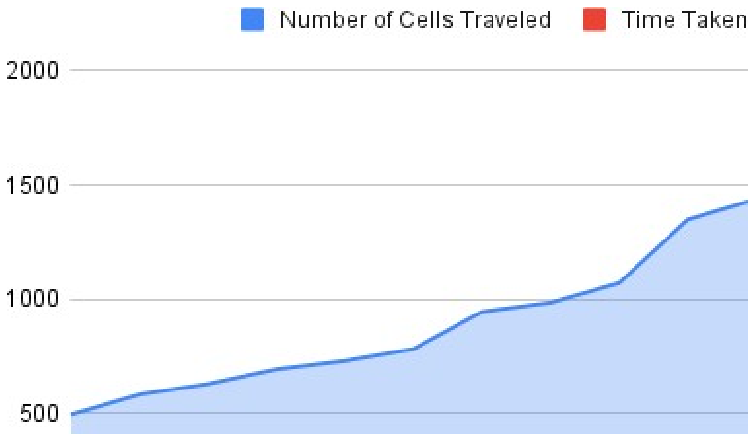

Figure 5: Original U-Net Performance on Training Data

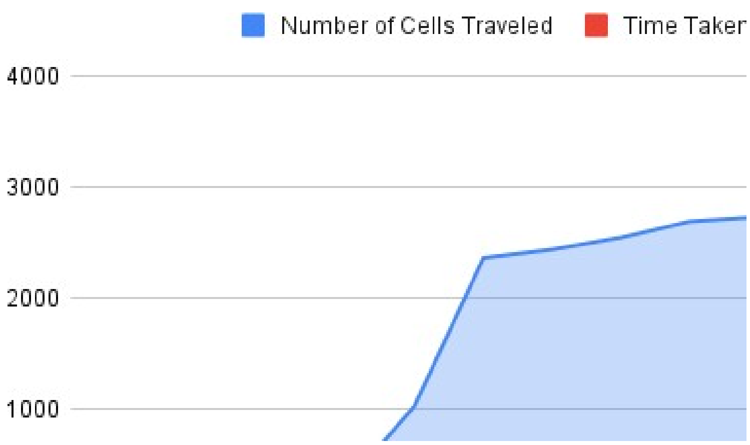

Figure 6: Original U-Net Performance on Test Data

4.3 Performance of the Modified U-Net

Our modified U-Net architecture, incorporating attention regions and opti- mized convolutional processes, showed further improvements:

- On the training dataset, it located the first survivor at just 69 unit time and all survivors at 1630 unit time.
- On unseen data, all survivors were located at 2928 unit time.
- Stress testing results:
  - Training set averages: First survivor at 69 unit time, all survivors at 1560 unit time.
  - Validation set averages: First survivor at 82 unit time, all sur- vivors at 2829 unit time.

The dramatic improvement in initial survivor detection time is particu- larly significant. In real-world SAR operations, rapid initial discoveries can boost team morale, provide crucial information about survivor

distribution, and inform broader strategic decisions.
Figures 7 and 8 show the performance of the modified U-Net on training and test data, respectively.

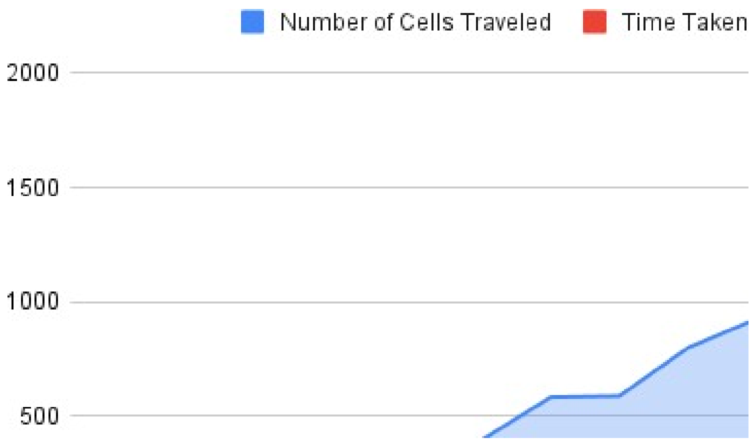

Figure 7: Modified U-Net Performance on Training Data

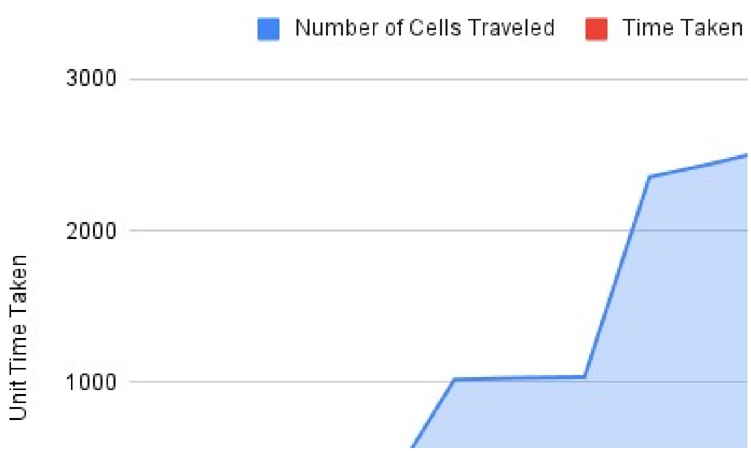

Figure 8: Modified U-Net Performance on Test Data

4.4 Comparative Analysis

To provide a clear overview of the improvements achieved, we present a comparative analysis of all approaches in Figure 9.
The modified U-Net consistently outperformed all other approaches across various metrics and scenarios. Key improvements include:

- Significantly reduced time to first survivor detection, crucial for real- world rescue operations.
- Lower overall search time, indicating more efficient path planning and resource utilization.
- Improved generalization to unseen data, suggesting better adaptability to new and unexpected scenarios.

We attribute these improvements to two key modifications in our ap- proach:

1. The introduction of attention regions, allowing the model to focus on areas of high importance or complexity within the search grid.

2. The removal of overlapping sliding kernels during convolution, reducing redundant computations and allowing for more efficient processing of the input data.

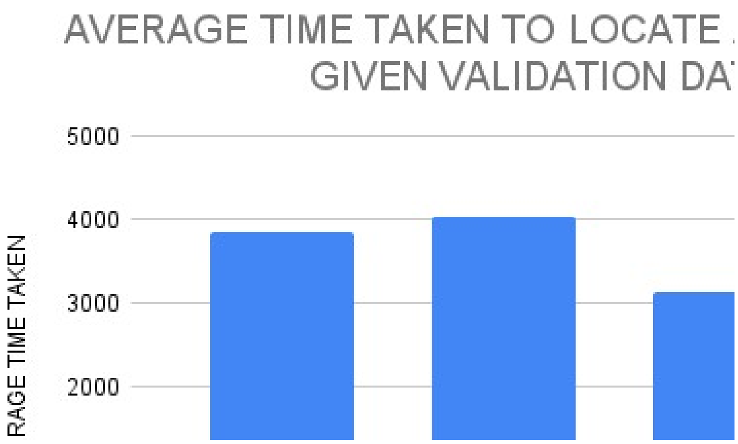

Figure 9: Comparative Analysis of All Approaches

These modifications make the model more suitable for the specific chal- lenges of SAR contexts, where rapid initial detections and efficient overall search patterns are crucial. The consistent performance improvements, particularly in the modified U-Net, showcase the adaptability and robustness of our approach.

4.5 Implications for Real-World SAR Operations

The results of our simulations have several important implications for real- world SAR operations:

- Rapid Initial Detection: The significant reduction in time to first sur- vivor detection could be life-saving in critical situations where every minute counts.

- Efficient Resource Utilization: By optimizing the search path, our ap- proach could allow SAR teams to cover more ground with fewer re- sources, potentially expanding the scope of rescue operations.

- Adaptability: The framework's ability to incorporate new information in real-time and adjust its strategy accordingly is crucial in the dynamic and unpredictable environments typical of disaster scenarios.

- Scalability: While our simulations focused on a specific grid size, the approach shows promise for scalability to larger and more complex search areas.

These findings suggest that our method could significantly enhance the efficiency of real-world SAR operations, leading to faster survivor recovery and improved mission outcomes. However, it is important to note that real-world implementation would require further testing and refinement to address challenges not captured in our simulations.

5 Conclusion

This research introduces a novel framework for path planning in robotic search and rescue (SAR) missions, centered on the generation of real-time amalgamated cost maps using a modified U-Net neural network architecture. Our approach represents a significant advancement in applying artificial intel- ligence and robotics to critical humanitarian efforts, offering a more adaptive and efficient solution to the complex challenges of SAR operations.

5.1 Key Contributions

The key aspects of our framework that contribute to its effectiveness include:

- Integration of static topological data with dynamic mission findings, allowing for a comprehensive understanding of the search environment.

- Development of an amalgamated cost map that prioritizes urgent and accessible regions, optimizing the balance between the importance of an area and the difficulty of reaching it.

- Implementation of a modified U-Net neural network architecture, specif- ically adapted for SAR applications, enabling adaptive cost prediction in complex, evolving environments.

- Continuous cost map updates based on real-time data collection, en- suring that the path planning remains responsive to new discoveries and changing conditions throughout the mission.

5.2 Performance Improvements

Our framework demonstrated promising results in extensive simulations, par- ticularly excelling in efficient survivor location. The modified U-Net model located survivors in just 1630 time units, a significant improvement over the 4024 units required by a naive approach and the 2413 units needed by a static cost map method. This performance underscores the potential of our framework to dramatically reduce response times in real-world crisis scenarios.

Notably, the rapid initial survivor detection capability of our system, with first detections occurring as early as 69 time units, could be particularly crucial in time-sensitive rescue operations where early successes can guide broader strategic decisions and resource allocation.

5.3 Limitations and Future Work

While our results are promising, we acknowledge certain limitations in our current study:

- Lack of standardized key performance indicators (KPIs) for quantita- tive evaluation of SAR robotics systems, making direct comparisons with other approaches challenging.

- Testing confined to simulated environments, which, while sophisticated, may not fully capture the complexity and unpredictability of real-world disaster scenarios (Wang et al., 2017).

These limitations point to several exciting directions for future research:

- Development of comprehensive and standardized KPIs for objective performance evaluation of SAR robotics systems across various scenar- ios and conditions.

- Implementation and rigorous testing of the framework on physical robotic platforms in diverse, real-world conditions to validate its practicality and effectiveness.

- Continuous refinement of the neural network architecture and training methodologies based on data collected from dynamic and unpredictable environments.

- Exploration of potential synergies with other advanced technologies, such as reinforcement learning (Niroui et al., 2019), to further optimize decision-making capabilities in complex SAR scenarios.

- Investigation of multi-robot coordination strategies using our frame- work, potentially enhancing the scalability and coverage of SAR operations.

5.4 Broader Implications

In conclusion, this work represents a significant step towards enhancing SAR operations through the integration of advanced AI and robotics technolo- gies. By bridging the gap between theoretical models and practical imple- mentation, our framework has the potential to revolutionize crisis response strategies.

The implications of this research extend beyond immediate improvements in SAR efficiency. As climate change potentially increases the frequency and severity of natural disasters, the need for advanced, AI-driven rescue technologies becomes ever more pressing. Our framework contributes to a growing body of research that aims to leverage cutting-edge technology to save lives and mitigate the impact of catastrophic events.

As we continue to refine and expand this approach, we aim to provide a robust, adaptable, and highly effective tool for emergency responders world- wide. The potential to save lives, reduce risks for human rescuers, and op- timize resource utilization in critical emergencies underscores the profound impact that AI and robotics can have in humanitarian applications.

Future work will focus on addressing the identified limitations, expanding the scope of our simulations, and moving towards real-world trials. We be- lieve that continued research in this direction will not only advance the field of SAR robotics but also contribute to broader discussions about the role of AI in disaster response and crisis management.

References

Applegate, D., Bixby, R., Chv´atal, V., & Cook, W. (2006, 01). The traveling salesman problem: A computational study. *The Traveling Salesman Problem: A Computational Study*.

Cormen, T., Leiserson, C., Rivest, R., & Stein, C. (2009). *Introduction to algorithms, third edition*. MIT Press.

Cully, A., Clune, J., Tarapore, D., & Mouret, J.-B. (2015, May). Robots that can adapt like animals. *Nature*, *521* (7553), 503–507. Retrieved from http://dx.doi.org/10.1038/nature14422 doi: 10.1038/na- ture14422

Gutin, G., Punnen, A., Barvinok, A., Gimadi, E., & Serdyukov, A. (2001, 06). The traveling salesman problem and its variations.

Held, M., & Karp, R. (1971, 01). The traveling salesman problem and minimum spanning trees. *Oper. Res.*, *18* . doi: 10.1287/opre.18.6.1138

Johnson, D. S., & McGeoch, L. A. (1997). The traveling salesman problem: a case study. *Local search in combinatorial optimization*, 215–310.

Kober, J., Bagnell, J., & Peters, J. (2013, 09). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, *32* , 1238-1274. doi: 10.1177/0278364913495721

Laporte, G. (1992, 06). The traveling salesman problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, *59* , 231-247. doi: 10.1016/0377-2217(92)90138-Y

Niroui, F., Zhang, K., Kashino, Z., & Nejat, G. (2019, 01). Deep reinforce- ment learning robot for search and rescue applications: Exploration in unknown cluttered environments. *IEEE Robotics and Automation Letters*, *PP* , 1-1. doi: 10.1109/LRA.2019.2891991

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *CoRR*, *abs/1505.04597* . Retrieved from http://arxiv.org/abs/1505.04597

Rosenkrantz, D. J., Stearns, R. E., & Lewis, P. M., II. (1977). An analysis of several heuristics for the traveling salesman problem. *SIAM Journal on Computing*, *6* (3), 563-581. doi: 10.1137/0206041

Wang, Z., Bapst, V., Heess, N., Mnih, V., Munos, R., Kavukcuoglu, K., & de Freitas, N. (2017). *Sample efficient actor-critic with experience replay.*

Zhu, Y., Mottaghi, R., Kolve, E., Lim, J. J., Gupta, A., Fei-Fei, L., & Farhadi, A. (2016). *Target-driven visual navigation in indoor scenes using deep reinforcement learning.*