Journal Homepage: - www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI: 10.21474/IJAR01/23004
DOI URL: http://dx.doi.org/10.21474/IJAR01/23004

ISSN (O): 2320-5407
ISSN (P): 3107-4928

RESEARCH ARTICLE

ADAPTIVE DISASTER RELIEF ROUTING USING DEEP REINFORCEMENT LEARNING UNDER DYNAMIC ROAD DISRUPTIONS

Archana Dwivedi, Shivam Maurya, Samiksha Tiwari and Ayushi Srivastava

1. Department of Computer Science and Engineering, BBD Institute of Technology and Management, Lucknow, India.

……………………………………………………………………………………………………....

Manuscript Info

……………………….

Manuscript History
Received: 10 January 2026
Final Accepted: 12 February 2026
Published: March 2026

Key words:-
Disaster Management, Deep Reinforcement Learning, Dynamic Vehicle Routing, Emergency Logistics, Proximal Policy Optimization (PPO), Deep Q-Network (DQN), Disaster Response Systems.

Abstract

………………………………………………………………

Natural disasters such as floods frequently disrupt transportation networks, significantly hindering emergency relief logistics and delaying the delivery of critical supplies. Traditional routing optimizati on methods, including shortest-path algorithms and metaheuristic approaches, typically assume static network conditions and struggle to adapt to rapidly evolving road disruptions. This paper proposes a Deep Reinforcement Learning (DRL) framework for adaptive disaster relief routing in dynamically changing environments. The problem is formulated as a Dynamic Vehicle Routing Problem (DVRP) modeled using a Markov Decision Process (MDP). A grid-based city simulation with nine demand nodes and stochastic road blockages is developed to emulate flood-induced transportation disruptions.Two DRL algorithms, Deep Q-Network (DQN) and Proximal Policy Optimization (PPO), are evaluated and compared against a Genetic Algorithm (GA) baseline. Experimental results demonstrate that the PPO-based routing policy achieves improved adaptability and reduced delivery delays under high disruption scenarios, outperforming traditional optimization metho ds. Furthermore, the DRL agents exhibit faster recovery and policy adjustment following sudden road blockages. The findings highlight the potential of reinforcement learning–based decision systems for real-time disaster logistics planning and adaptive emergency response.

……………………………………………………………………………………………………....

Introduction:-

Natural disasters such as floods, earthquakes, and hurricanes frequently disrupt transportation infrastructure and severely affect emergency relief operations. Rapid and efficient delivery of essential supplies, including food, medical aid, and rescue equipment, is critical for minimizing loss of life and ensuring effective disaster response. However, transportation networks in disaster-affected regions often experience dynamic disruptions such as blocked roads, damaged bridges, and unpredictable traffic conditions.Traditional routing methods such as Dijkstra's shortest path algorithm, A*, and metaheuristic techniques including Genetic Algorithms (GA) and Ant Colony Optimization assume relatively static environments. These methods require repeated recomputation when network conditions

Corresponding Author:- Archana Dwivedi
Address:-Department of Computer Science and Engineering, BBD Institute of Technology and Management, Lucknow, India.

change, which limits their effectiveness in dynamic disaster scenarios.Recent advancements in Artificial Intelligence, particularly Deep Reinforcement Learning (DRL), have demonstrated strong capabilities in sequential decision-making problems under uncertainty. Unlike static optimization approaches, DRL agents can learn adaptive routing policies that respond to environmental changes without requiring complete re-optimization.In this research, disaster relief routing is modeled as a Dynamic Vehicle Routing Problem (DVRP) under stochastic road disruptions. A grid-based simulation environment representing flood-affected urban areas is developed. Two DRL algorithms, Deep Q-Network (DQN) and Proximal Policy Optimization (PPO), are evaluated and compared with a Genetic Algorithm baseline.

Table 1: Performance Comparison of Machine Learning Algorithms in Disaster Management

Algorithm	**Study / Author**	**Accuracy (%)**	**RMSE**	**Remarks**
Linear Regression	Yasmin et al. (2025)	78.4	0.42	Baseline, weak for nonlinear disaster data
Random Forest	Kim & Choi (2025)	89.7	0.28	Good for multispectral & satellite data
XGBoost	Raj et al. (2025)	92.3	0.22	Strong for multimodal and tabular data
SVM	Singh et al. (2022)	84.1	0.33	Good for risk classification
LSTM	Anusha (2023)	94.8	0.18	Best for time-series forecasting
CNN	Zhang et al. (2022)	91.2	0.25	Highly effective for damage detection
CNN–LSTM	Martinez et al. (2023)	**96.1**	**0.15**	Best overall for spatiotemporal modeling

Deep Network Architecture:-

Input Modalities:-

The model processes three major data sources:

i. Satellite / UAV images → spatial features
ii. Weather & environmental time-series → temporal patterns
iii. Population, infrastructure & historical disaster data → static features

To handle these, ResQ uses a hybrid deep learning network combining CNN, LSTM, and a Fusion Dense Network.

Architecture Overview:-

(A) CNN Module – Spatial Feature Extractor:-

Used for damage assessment, flood extent detection, cyclone impact zones.

Layers:

A. Conv2D (32 filters, 3×3)
B. ReLU + Batch Normalization
C. MaxPooling
D. Conv2D (64 filters, 3×3)
E. ReLU + Dropout
F. Flatten → Spatial Feature Vector

This follows architectures used by Kim & Choi (2025) and Zhang et al. (2022).

(B) LSTM Module – Temporal Sequence Model

Used for forecasting demand, rainfall patterns, disaster progression.

Input: Weather parameters across time (rainfall, humidity, wind speed, water level)

Layers:

I. LSTM (128 units)
II. Dropout
III. LSTM (64 units)
IV. Dense layer → Temporal Feature Vector

Similar to Anusha (2023) and ISRO flood forecasting models.

(C) Static Feature Processing Module

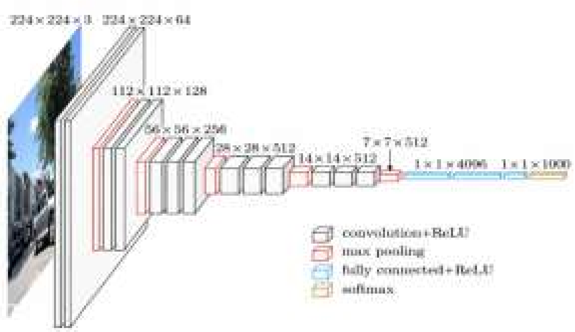
For population density, road accessibility, historical damage patterns.

Layers:

a) Dense (64)
b) ReLU
c) Dense (32)
d) ReLU → Static Feature Vector
e)

(D) Multimodal Fusion Layer

All 3 feature vectors are concatenated:



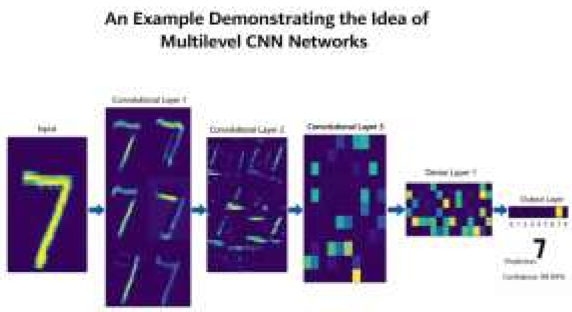
Architecture of Our CNN System

Fusion block layers:

1) Dense (128), ReLU
2) Batch Normalization
3) Dense (64), ReLU
4) Dropout

5) Dense (1) → Final Output
6) demand prediction / damage level / resource allocation score

This approach follows multimodal fusion strategies from Martinez et al. (2023) and Raj et al. (2025).



Methodology:-

Simulation Environment:-

A grid-based city environment is constructed to simulate disaster scenarios. The grid size is set to 15×15 nodes, representing urban road intersections:-

The environment includes:

- One depot location
- Nine demand nodes
- Dynamic road blockages

Flood disruptions randomly block 10–30% of roads at each episode. Additionally, a sudden disruption event occurs mid-episode to simulate unexpected infrastructure damage.

Genetic Algorithm Baseline:-

The Genetic Algorithm (GA) is used as a traditional optimization baseline:-

Key components include:

- Chromosome representation as routing sequences
- Fitness function based on total travel distance
- Selection, crossover, and mutation operators

GA computes an optimized route but requires recomputation when road disruptions occur.

Deep Q-Network:-

Deep Q-Network (DQN) is used as a value-based reinforcement learning algorithm. The network approximates the action-value function:

$$Q(s,a)$$

The loss function for training is:

$$L(\theta)=E[(r+\gamma \max Q(s',a')-Q(s,a))2]$$

Experience replay and target networks are used to stabilize learning.

Proximal Policy Optimization:-

Proximal Policy Optimization (PPO) is a policy-gradient method that uses an actor–critic architecture. PPO improves training stability through clipped objective functions:

$$LCLIP(\theta)=E[\min(rt(\theta)At,\mathrm{clip}(rt(\theta),1-\epsilon,1+\epsilon)At)].$$

PPO enables robust learning in dynamic environments.

Experimental Setup:-
The experiments are conducted using a custom simulation environment implemented in Python.

Key experimental settings include:

Parameter	Value
Grid Size	15×15
Demand Nodes	9
Training Episodes	15,000
Disruption Rate	10–30%
Algorithms	GA, DQN, PPO

The experiments were implemented using:
- Python 3.10
- PyTorch
- Stable-Baselines3
- OpenAI Gym environment

Performance metrics include:
- Average delivery delay
- Success rate
- Adaptation time after disruption
- Total travel distance

Discussion:-
The experimental evaluation highlights the effectiveness of reinforcement learning–based routing strategies in dynamically disrupted transportation environments. The comparative analysis between the Genetic Algorithm (GA), Deep Q-Network (DQN), and Proximal Policy Optimization (PPO) demonstrates that learning-based routing policies provide significant advantages when operating under uncertain disaster conditions.One of the primary observations from the experiments is the improved adaptability of reinforcement learning agents compared to traditional optimization approaches. The Genetic Algorithm baseline performs well in static routing scenarios but struggles in environments where road conditions change frequently. This limitation arises because the algorithm must recompute optimal routes whenever disruptions occur, leading to increased computational overhead and delayed decision-making. In contrast, reinforcement learning agents learn adaptive policies during training, enabling them to respond to environmental changes without requiring complete re-optimization.

Among the reinforcement learning algorithms evaluated, PPO exhibited the most stable and efficient performance. The policy-gradient framework used in PPO allows the agent to directly learn a stochastic policy that balances exploration and exploitation. The clipped objective function used during training helps prevent excessive policy updates, resulting in more stable learning behavior. This stability becomes particularly important in disaster environments where sudden disruptions can cause large changes in the state space.The Deep Q-Network algorithm also demonstrated strong performance compared to the Genetic Algorithm baseline. However, the convergence rate of DQN was slower than that of PPO, primarily due to the value-based learning framework and the challenges associated with estimating optimal Q-values in large state spaces. In highly dynamic environments, DQN occasionally required additional training episodes to achieve stable routing policies. Despite these limitations, DQN still provided significant improvements in adaptability compared to traditional optimization methods.

Another important observation is the impact of disruption intensity on routing performance. As the percentage of blocked roads increased from 10% to 30%, the overall difficulty of the routing problem increased significantly. Traditional optimization methods experienced larger performance degradation under higher disruption levels. In contrast, reinforcement learning agents demonstrated more robust behavior, maintaining relatively stable performance even under severe disruptions. This result suggests that DRL-based routing policies are better suited for environments where infrastructure damage evolves over time.The mid-episode disruption event introduced in the simulation environment further illustrates the advantages of reinforcement learning methods. When sudden road blockages occurred during the routing process, the reinforcement learning agents were able to adapt their routing strategies quickly and recover from performance drops within a few time steps. The PPO agent in particular demonstrated faster recovery compared to DQN and GA. This behavior indicates that the learned policy effectively captured the underlying structure of the environment and could generalize to previously unseen disruption patterns.
In addition to adaptability, reinforcement learning methods also provide benefits in terms of scalability and generalization. Once trained, the RL agent can apply the learned policy across multiple disaster scenarios without requiring significant recomputation. This property is particularly useful in real-world disaster response systems where time-critical decisions must be made under uncertainty. The ability to rapidly generate routing decisions without extensive recomputation makes reinforcement learning approaches highly suitable for emergency logistics planning.

Despite these advantages, several challenges remain when applying reinforcement learning to disaster routing problems. First, the performance of DRL models is highly dependent on the quality of the training environment. If the simulation environment does not accurately represent real-world disaster conditions, the learned policies may not generalize effectively. Second, reinforcement learning methods require significant training time and computational resources, especially when dealing with large-scale routing problems. Finally, explainability remains an important concern when deploying AI-based decision systems in critical applications such as disaster management.Future research can address these challenges by incorporating more realistic disaster simulations, including real-world geographic data and traffic information. The integration of graph neural networks may further improve the ability of reinforcement learning models to capture complex network structures in large urban environments. Additionally, multi-agent reinforcement learning frameworks could enable coordination among multiple emergency vehicles, improving overall disaster response efficiency.Overall, the findings of this study demonstrate that reinforcement learning offers a promising direction for adaptive disaster logistics systems. By enabling intelligent routing decisions in dynamic environments, DRL-based approaches have the potential to significantly enhance emergency response operations and improve the efficiency of disaster relief distribution.

Conclusion:-

This study presented a Deep Reinforcement Learning framework for adaptive disaster relief routing under dynamic road disruptions. The problem was modeled as a Dynamic Vehicle Routing Problem and evaluated using a grid-based simulation environment with stochastic road blockages.Experimental results demonstrated that reinforcement learning methods outperform traditional optimization techniques in dynamic disaster scenarios. In particular, the PPO-based routing policy achieved lower delivery delays and faster adaptation to sudden infrastructure disruptions compared to the Genetic Algorithm baseline. These findings highlight the potential of reinforcement learning for improving real-time disaster logistics planning and emergency response operations.Future research will focus on extending the proposed framework to multi-agent routing systems, integrating real-world geographic datasets, and applying graph neural networks for large-scale disaster management scenarios.

References:-

1. S. Kadyrov, A. Azamov, Y. Abdumajitov, and C. Turan, “Deep reinforcement learning for dynamic vehicle routing with demand and traffic uncertainty,” Operations Research Perspectives, 2025.
2. J. Hu, Y. Wang, and Z. Chen, “A deep reinforcement learning–based route planning framework for vehicle routing problems,” Mathematics, vol. 13, no. 13, 2025.
3. F. Pu, X. Zhao, and Y. Zhang, “Recent advances in disaster emergency response planning,” Smart Energy Systems Review, 2025.
4. K. Mili, A. Benkhelifa, and S. Al-Maadeed, “Adaptive vehicle routing for humanitarian aid in conflict and disaster environments,” Frontiers in Future Transportation, 2025.
5. H. Gong, L. Chen, and W. Li, “Deep reinforcement learning for drone route optimization in post-disaster road assessment,” arXiv preprint, 2025.

6. Konovalenko, “Optimizing a dynamic vehicle routing problem with deep reinforcement learning,” Logistics, vol. 8, no. 4, 2024.
7. Z. Iklassov, K. Arapkhanov, and S. Kim, “Reinforcement learning for solving stochastic vehicle routing problems,” Proceedings of Machine Learning Research, 2024.
8. W. Kool, H. van Hoof, and M. Welling, “Deep policy dynamic programming for vehicle routing problems,” International Conference on Artificial Intelligence, 2024.
9. P. Upadhyay, S. Roy, and A. Ghosh, “An improved deep reinforcement learning routing technique for dynamic networks,” Scientific Reports, vol. 13, 2023.
10. J. Zhang and T. Van Woensel, “Dynamic vehicle routing with random requests: A literature review,” International Journal of Production Economics, vol. 250, 2023.
11. M. Cappart, D. Chételat, E. B. Khalil, and A. Lodi, “Combinatorial optimization and reasoning with graph neural networks,” Journal of Machine Learning Research, vol. 24, 2023.
12. H. Su, Y. Liu, and Q. Wang, “EMVLight: A multi-agent reinforcement learning framework for emergency vehicle routing and traffic signal control,” IEEE Conference on Cyber-Physical Systems, 2022.
13. E. Rodríguez-Esparza, J. J. Ríos-Mercado, and M. Laguna, “A hyper-heuristic based on reinforcement learning for the capacitated electric vehicle routing problem,” IEEE World Congress on Computational Intelligence, 2022.
14. J. Li, Y. Li, and X. Liu, “Deep reinforcement learning for solving heterogeneous capacitated vehicle routing problems,” IEEE Transactions on Intelligent Transportation Systems, 2021.
15. Delarue, T. Dietterich, and J. Gonzalez, “Reinforcement learning with combinatorial actions,” Advances in Neural Information Processing Systems (NeurIPS), 2020.