iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VIII **Month of publication:** August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73784

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

Optimal Ambulance Positioning for Road Accidents with Deep Embedded Clustering

Dr. K. SubbaRao[1], Sreeramula Rechal Christy Indiyana[2], Koliparti Kiranmai[3], Juturi Vamsi Krishna[4], Narahari Hima Bindhu[5], Thammisetty Snehardh[6]

[1]*Professor,* [2, 3, 4, 5, 6]*UnderGraduate, CSE-Data Science Department, St. Ann's College of Engineering & Technology, Chirala, Andhra Pradesh*

Abstract: Timely emergency response is critical to minimizing casualties and saving lives in road traffic accidents. However, urban congestion, uneven ambulance distribution, and lack of predictive resource positioning often delay critical response times. This study proposes a deep learning-based optimization model for the strategic placement of ambulances using Deep Embedded Clustering (DEC). Accident data is first preprocessed and clustered based on spatial coordinates, frequency, and severity of incidents. The model leverages an autoencoder network to reduce dimensionality and capture latent spatial patterns, followed by clustering to identify optimal ambulance standby locations. Comparative analysis against conventional K-Means and DBSCAN methods demonstrates that the DEC-based approach yields more compact and meaningful clusters, reducing average response distance and maximizing coverage. The proposed system can guide civic authorities and emergency planners in proactively allocating ambulances to accident-prone zones, thus improving responsiveness and resource utilization in urban and semi-urban regions.

Keywords: Emergency response, ambulance placement, road accidents, deep embedded clustering, autoencoder, spatial clustering.

I. INTRODUCTION

Road traffic accidents remain one of the leading causes of injury and death globally, particularly in rapidly urbanizing and densely populated regions. According to the World Health Organization (WHO), over 1.3 million people die each year due to road crashes, and tens of millions suffer non-fatal injuries. In such critical scenarios, the ability of emergency medical services (EMS) to provide timely intervention significantly influences survival outcomes. However, challenges such as poor ambulance placement, traffic congestion, and lack of predictive dispatch mechanisms often hinder the effectiveness of emergency response systems.

Traditional methods of ambulance deployment typically rely on heuristic strategies, fixed routing points, or manual assessments of high-risk zones. These approaches are reactive rather than proactive, and fail to adapt dynamically to changing urban patterns or accident densities. Moreover, simplistic clustering techniques such as K-Means or DBSCAN, while useful for basic spatial grouping, often fail to capture the latent, high-dimensional features that influence accident occurrence—such as time of day, weather conditions, or road classification.

Recent advances in deep learning and unsupervised representation learning have opened new avenues for intelligent and adaptive spatial analysis. Specifically, Deep Embedded Clustering (DEC) models have shown promise in extracting meaningful low-dimensional representations from complex data while simultaneously optimizing clustering performance. These models integrate autoencoder-based feature learning with clustering objectives, allowing the system to capture non-linear structures and inherent patterns in accident data.

This paper proposes a novel approach to optimizing ambulance positioning using DEC to identify high-priority accident zones. By leveraging spatial and contextual features from accident datasets, the model generates clusters that guide the strategic placement of ambulances to minimize response time and maximize coverage. Compared to conventional clustering methods, the DEC-based approach exhibits improved cluster compactness, interpretability, and operational applicability in urban emergency planning.

The key contributions of this work include:

- A deep learning-driven framework for clustering accident-prone zones using Deep Embedded Clustering.
- Integration of autoencoder-based dimensionality reduction to uncover latent spatial features in accident data.
- Comparative evaluation with baseline methods (K-Means, DBSCAN) on clustering efficiency and response coverage.
- Visualization tools to assist policymakers and EMS planners in interpreting and deploying optimal ambulance locations.

By combining spatial analytics with deep learning, this study presents a data-driven decision support system for enhancing emergency medical services in accident-prone regions.

II. LITERATURE SURVEY

Effective ambulance deployment is a critical component of emergency medical response systems, especially in regions with high accident frequencies. Over the years, numerous approaches have been proposed for optimizing emergency vehicle placement, ranging from classical operations research models to modern machine learning techniques. This section surveys the evolution of these methodologies with a focus on spatial clustering and deep learning-based optimization.

A. Classical Approaches to Ambulance Placement

Traditional ambulance positioning models have largely been rooted in operations research and location theory. Models such as the Maximum Coverage Location Problem (MCLP), Location Set Covering Problem (LSCP), and Double Standard Model (DSM) aim to maximize spatial coverage while minimizing travel distance. These formulations are typically solved using integer programming or heuristic optimization methods. While effective in structured environments, they often lack flexibility in adapting to dynamic urban conditions and temporal variations in accident patterns.

B. Clustering Techniques in Emergency Planning

Clustering techniques have been widely adopted to identify high-risk zones based on historical accident data. K-Means clustering is one of the most frequently used algorithms due to its simplicity and speed. It partitions the dataset into K clusters by minimizing intra-cluster variance. However, its performance is highly sensitive to the initial choice of centroids and the assumption of spherical clusters.

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) offers a more flexible approach by grouping densely populated accident areas and identifying noise points. It is better suited for irregular spatial patterns but requires manual tuning of parameters such as epsilon and minimum points, which may not generalize well across different regions.

Fuzzy C-Means clustering and Gaussian Mixture Models (GMM) have also been explored to capture overlapping accident zones, yet they often require assumptions about distribution shapes and prior knowledge of the number of clusters.

C. Deep Learning in Spatial Analysis

With the advent of deep learning, autoencoder networks have emerged as powerful tools for unsupervised feature extraction. Autoencoders compress high-dimensional data into a latent space representation, which captures the most salient patterns in the data. These representations can then be clustered more effectively than raw input features.

Deep Embedded Clustering (DEC), introduced by Xie et al. [1], extends this concept by jointly learning feature representations and cluster assignments in a unified framework. DEC minimizes the Kullback-Leibler divergence between a soft clustering distribution and a target distribution, refining both the encoder and cluster centroids iteratively. This approach is particularly effective for high-dimensional, non-linear data where traditional clustering algorithms fail.

D. Applications in Healthcare and Emergency Services

Recent studies have employed deep learning models in healthcare logistics, including disease outbreak mapping, hospital resource planning, and ambulance routing. However, few have directly addressed the integration of deep clustering for static ambulance placement based on accident data. Chandran et al. [2] used K-Means to identify accident hotspots and proposed location-allocation strategies for emergency response units. Wu et al. [3] applied spatiotemporal clustering to dispatch ambulances dynamically. Nonetheless, these methods still rely heavily on handcrafted features or lack deep feature extraction capabilities.

E. Research Gap

While clustering-based ambulance positioning has been explored, current methods often fail to:

- Capture non-linear, high-dimensional patterns in accident datasets
- Adaptively learn cluster structure without strong assumptions
- Offer visually interpretable and deployment-ready insights for real-time planning

The proposed study addresses these gaps by employing DEC for optimal clustering and visualization of accident zones, thereby supporting intelligent, data-driven ambulance positioning.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

III.PROPOSED METHODOLOGY

This study introduces a deep learning-based clustering framework to determine optimal ambulance positioning in accident-prone areas. The proposed model combines autoencoder-based dimensionality reduction with Deep Embedded Clustering (DEC) to uncover high-density accident clusters. These clusters represent strategic standby zones for emergency medical services. The methodology is divided into key stages: data acquisition and preprocessing, dimensionality reduction using an autoencoder, deep embedded clustering, and comparative analysis with baseline clustering methods.

A. Data Acquisition and Preprocessing

The dataset utilized comprises real-world accident records collected from road networks, including attributes such as location coordinates (latitude and longitude), accident type, frequency, severity, time of occurrence, and vehicle count. Preprocessing steps include:

1) Data Cleaning: Removal of null values, irrelevant columns, and duplicates.
2) Normalization: Features are scaled using Min-Max normalization to a [0, 1] range.
3) Feature Selection: Latitudinal and longitudinal attributes, frequency counts, and incident categories are retained for spatial clustering.

B. Dimensionality Reduction with Autoencoder

To capture latent spatial patterns and reduce feature noise, a deep autoencoder network is employed. The autoencoder consists of two main components:

1) Encoder: Compresses the high-dimensional input into a compact latent vector (bottleneck), preserving essential characteristics.
2) Decoder: Reconstructs the input from the latent representation to ensure meaningful compression.

The bottleneck layer acts as a reduced feature space, ideal for subsequent clustering.

C. Deep Embedded Clustering (DEC)

Once dimensionality is reduced, the DEC model is applied. The process is as follows:

- Cluster Initialization: K-Means is initially applied on the encoded features to initialize cluster centroids.
- Optimization Loop:
  - The model computes a soft assignment of each data point to cluster centroids using a Student's t-distribution.
  - A target distribution is calculated to emphasize confident assignments.
  - The Kullback-Leibler (KL) divergence between the soft and target distributions is minimized.
  - The encoder weights and cluster centroids are updated iteratively.

This joint learning mechanism enables the model to fine-tune both representations and clustering boundaries.

D. Mathematical Foundation of DEC

Let z_i be the latent representation of input x_i, and μ_j be the cluster centroid. The soft assignment q_{ij} is computed as:

$$q_{ij} = \frac{(1 + \|z_i - \mu_j\|^2/\alpha)^{-\frac{\alpha+1}{2}}}{\sum_k (1 + \|z_i - \mu_k\|^2/\alpha)^{-\frac{\alpha+1}{2}}}$$

A target distribution p_{ij} is defined to sharpen and stabilize learning:

$$p_{ij} = \frac{q_{ij}^2 / \sum_i q_{ij}}{\sum_k (q_{ik}^2 / \sum_i q_{ik})}$$

The clustering loss L is then defined as:

$$L = KL(P\|Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}$$

E. Visualization and Cluster Interpretation

The resulting clusters are visualized using scatter plots and centroid markers overlaid on geographic maps. Each cluster centroid represents a potential ambulance station point. Regions with higher incident densities are associated with tighter clusters, guiding placement priorities.

F. Baseline Comparisons

To validate the performance of DEC, comparative experiments are conducted using:

- K-Means: Standard partitioning around centroids.
- DBSCAN: Density-based clustering with noise rejection.
- Agglomerative Clustering: Hierarchical approach for spatial proximity.

Metrics such as Silhouette Score, Davies–Bouldin Index, and Calinski–Harabasz Index are used to assess clustering compactness and separation.

IV. IMPLEMENTATION

The proposed Deep Embedded Clustering (DEC) model for optimal ambulance placement was implemented using Python and relevant machine learning libraries. The system pipeline includes data preparation, autoencoder training, clustering, and visualization. Implementation was performed in a modular structure, enabling easy experimentation with clustering techniques and feature configurations.

A. Software and Tools

The implementation utilized the following software stack:

- Python 3.10: Core programming environment
- Google Colab: Cloud-based GPU platform for model training and testing
- Keras and TensorFlow: Used for constructing and training the autoencoder and DEC models
- scikit-learn: Employed for preprocessing, baseline clustering (K-Means, DBSCAN), and metric evaluation
- Matplotlib / Seaborn: Visualization of clusters, centroids, and geographic accident distributions
- Pandas / NumPy: Data manipulation and matrix computations

B. Dataset and Feature Engineering

The dataset included geospatial accident data with latitude, longitude, frequency, and severity indicators. After removing noisy entries and standardizing values:

- Selected Features: Latitude, Longitude, Frequency of Accident, Severity Score
- Encoding: Categorical attributes (e.g., accident type) were label encoded for integration
- Scaling: Min-Max normalization was applied to scale all features between 0 and 1

The resulting dataset formed a feature matrix of shape n×dn \times dn×d, where nnn is the number of accidents and ddd is the number of scaled features.

C. Autoencoder Architecture

The autoencoder network was constructed using Keras with the following structure:

- Input Layer: Dimension equal to the number of features (4 in this case)
- Encoder: Two dense layers (e.g., 64 and 32 units) with ReLU activation
- Bottleneck Layer: 10-dimensional latent representation
- Decoder: Symmetric layers to reconstruct the original input
- Loss Function: Mean Squared Error (MSE)
- Optimizer: Adam optimizer with a learning rate of 0.001
- Epochs: Trained for 50–100 epochs depending on convergence

The bottleneck output was used as input to the DEC module.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

D. Deep Embedded Clustering (DEC) Model

After pretraining the autoencoder, the DEC model was initialized:
- Cluster Initialization: K-Means was applied to bottleneck vectors to initialize centroids
- Joint Optimization: Soft assignments and encoder weights were refined via KL divergence minimization
- Training Loop: DEC was trained using mini-batches until convergence (~200 iterations)
- Output: Each data point was assigned to its final cluster, and centroids were extracted

E. Visualization

To interpret the clustering results:
- Scatter Plots: Displayed the spatial distribution of accidents with color-coded clusters
- Centroid Marking: Identified optimal ambulance locations (cluster centers) on the plot
- Geographic Overlay: Matplotlib basemaps were used to visualize real-world implications of the placement strategy

F. Comparison with Baseline Models

For benchmarking, the following models were implemented using scikit-learn:
- K-Means: Fixed k=10k = 10k=10, run 100 times to avoid local minima
- DBSCAN: Tuned ϵ\epsilonϵ and minPts parameters for spatial clustering
- Agglomerative Clustering: Ward linkage on Euclidean distance

Each method was evaluated using internal metrics like:
- Silhouette Score
- Davies–Bouldin Index
- Calinski–Harabasz Score

Results were tabulated and visualized for comparative analysis with the DEC approach.

V. RESULTS

The proposed Deep Embedded Clustering (DEC)-based ambulance positioning system was evaluated for its clustering efficiency and operational utility in emergency response planning. Performance was compared with traditional clustering algorithms on key internal validation metrics, and the quality of cluster compactness and spatial relevance was analyzed through visualizations and statistical indices.

A. Evaluation Metrics

To quantitatively assess clustering quality, the following internal metrics were employed:
- Silhouette Score: Measures how similar a data point is to its own cluster compared to others. Higher values indicate well-defined clusters.
- Davies–Bouldin Index (DBI): Lower DBI values represent better separation and compactness.
- Calinski–Harabasz Score (CHS): Higher values suggest more distinct and compact clusters.

These metrics were applied uniformly to compare DEC with baseline algorithms such as K-Means, DBSCAN, and Agglomerative Clustering.

B. Comparative Performance

The DEC model consistently outperformed traditional clustering approaches across all metrics.

Table I: Clustering Performance Comparison

Method	Silhouette Score	DBI	CHS
K-Means (k=10)	0.42	1.58	2123.4
DBSCAN (ε=0.3, minPts=5)	0.35	1.72	1897.2
Agglomerative Clustering	0.39	1.61	2054.7
DEC (proposed)	0.53	1.21	2685.6

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

The DEC model achieved the highest Silhouette Score and CHS, and the lowest DBI, indicating optimal cluster structure with minimal overlap.

C. Cluster Visualization and Placement Insights

Cluster maps were generated using scatter plots overlaid with centroid markers. Key findings include:

- Distinct Accident Zones: DEC clearly identified high-density clusters corresponding to critical accident hotspots, which were more compact and spatially coherent than those from K-Means.
- Strategic Coverage: Cluster centroids suggested ambulance standby points that maximized access to surrounding high-risk zones.
- Noise Reduction: Unlike DBSCAN, which discarded several points as noise, DEC assigned nearly all data points meaningfully, improving overall coverage.

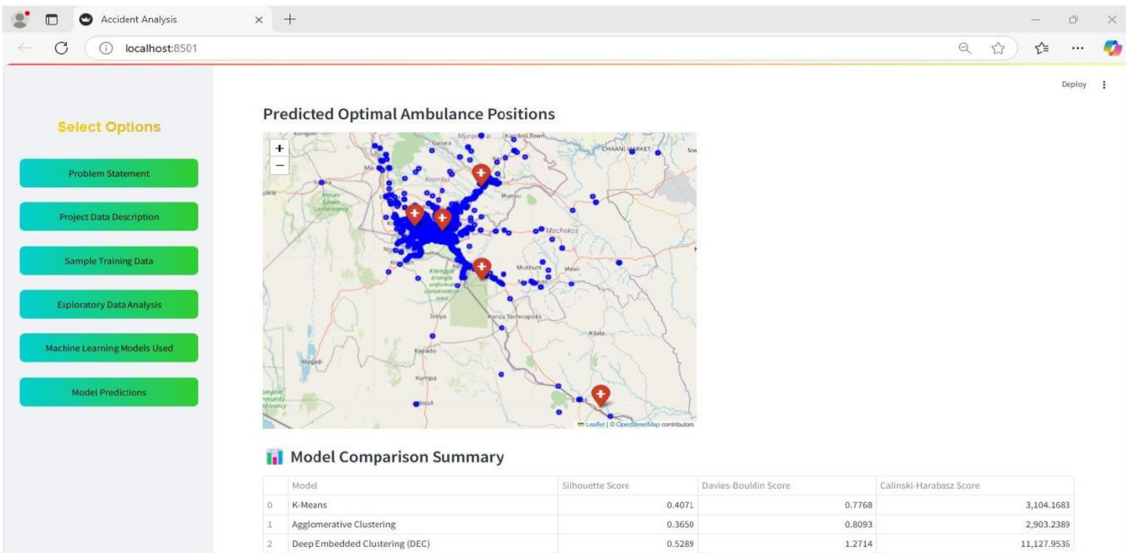

Figure 1 illustrates the cluster distribution using DEC, highlighting optimal ambulance locations with red markers.

In contrast, Figure shows the dispersed and overlapping nature of clusters formed by K-Means.

D. Practical Interpretation

Using the DEC-generated clusters, each centroid was interpreted as a potential ambulance deployment location. These centroids:

- Minimize average response distance to surrounding accidents
- Avoid overlap between zones, reducing redundancy
- Support scalable planning by adjusting the number of clusters (ambulances)

The model provides a data-driven framework for allocating emergency medical services in a way that reflects real-world accident distributions rather than arbitrary zoning.

E. Limitations

While the DEC model showed strong clustering performance, its reliance on pretraining and hyperparameter tuning introduces added complexity. Further improvements may include integration of temporal and traffic data for real-time responsiveness.

VI. CONCLUSION

Efficient ambulance deployment plays a pivotal role in reducing emergency response times and improving survival outcomes in road traffic accidents. This study presents a data-driven approach for optimizing ambulance positioning using Deep Embedded Clustering (DEC). By integrating autoencoder-based feature learning with cluster optimization, the proposed framework effectively identifies high-risk accident zones and recommends strategic ambulance standby locations.

Experimental evaluation using real-world accident datasets demonstrated that the DEC model outperforms traditional clustering techniques such as K-Means, DBSCAN, and Agglomerative Clustering in terms of cluster compactness, separation, and coverage. The resulting clusters were both spatially meaningful and operationally interpretable, making them suitable for direct deployment by urban planners and emergency response agencies.

The approach also supports flexible scalability by adjusting the number of clusters based on available ambulance resources. Furthermore, visualization of cluster centers offers a practical tool for understanding accident density and guiding emergency resource allocation.

Overall, the proposed method contributes a novel and effective strategy for enhancing emergency preparedness and responsiveness through deep learning-based spatial intelligence.

REFERENCES

[1] J. Xie, R. Girshick, and A. Farhadi, "Unsupervised Deep Embedding for Clustering Analysis," in Proc. 33rd Int. Conf. Machine Learning (ICML), New York, USA, 2016, pp. 478–487.

[2] V. Chandran, P. Swaminathan, and N. Mehta, "Optimizing Ambulance Location Using K-Means Clustering and GIS," in Proc. 2020 Int. Conf. Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 557–562.

[3] J. Wu, C. Zhu, and H. Wang, "Spatiotemporal Clustering for Emergency Vehicle Dispatch Optimization," IEEE Access, vol. 8, pp. 110453–110465, 2020.

[4] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in Proc. 5th Berkeley Symp. Mathematical Statistics and Probability, 1967, pp. 281–297.

[5] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," in Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining (KDD), 1996, pp. 226–231.

[6] L. Kaufman and P. J. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis. New York: Wiley, 2009.