Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

DEEP EMBEDDED CLUSTERING FOR OPTIMAL AMBULANCE POSITIONING IN RESPONSE TO ROAD ACCIDENTS

[1] *Umar Basha Shaik*, [2] *Dr.M.Subba Reddy*
[1] *M.Tech Student*, [2] *Associate Professor*
Department of Computer Science Engineering
SVR Engineering College, Nandyal

ABSTRACT

Efficient ambulance positioning is crucial for minimizing response times and improving emergency medical services (EMS) in road accident scenarios. Traditional ambulance deployment strategies often rely on static stationing models or historical data-based heuristics, which may not effectively adapt to dynamic accident patterns. This paper proposes a Deep Embedded Clustering (DEC)-based approach for optimal ambulance positioning, leveraging real-time accident data, traffic conditions, and geographic information. The model integrates deep learning-based feature extraction with unsupervised clustering techniques to dynamically identify high-risk accident zones and determine optimal ambulance locations. By embedding accident-related features into a lower-dimensional latent space, the DEC model enhances the clustering process, improving both spatial efficiency and response accuracy. Experimental evaluations using real-world accident datasets demonstrate that the proposed approach outperforms traditional k-means and density-based clustering models in terms of cluster stability, ambulance coverage efficiency, and response time reduction. This research highlights the potential of deep learning-driven optimization for EMS deployment, contributing to enhanced public safety and emergency response management.

Keywords: Deep Embedded Clustering, Emergency Medical Services, Ambulance Positioning, Road Accident Response, Deep Learning, Unsupervised Clustering.

I. INTRODUCTION

Road accidents are a major global public safety concern, often leading to severe injuries and fatalities. The efficiency of Emergency Medical Services (EMS) plays a crucial role in reducing casualty rates by ensuring timely medical intervention. However, traditional ambulance deployment strategies, which rely on static stationing models or historical accident data, often fail to account for dynamic traffic conditions, evolving accident patterns, and real-time emergency demands. These limitations result in delayed response times, potentially worsening patient outcomes. Therefore, optimizing ambulance positioning is essential to improving emergency response effectiveness.

Recent advancements in machine learning and deep learning have enabled data-driven solutions for intelligent ambulance placement. While conventional clustering techniques such as k-means and DBSCAN have been used to identify accident-prone locations, they struggle with high-dimensional accident data and dynamic environmental factors. To address these

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

challenges, this paper proposes a Deep Embedded Clustering (DEC)-based approach for optimal ambulance positioning in response to road accidents. The DEC model leverages deep learning for feature extraction and dimensionality reduction, followed by an unsupervised clustering process to identify high-risk accident zones and allocate ambulances accordingly.

The key contributions of this research include:

- Integration of deep learning with clustering to improve the accuracy of accident-prone area identification.
- A dynamic ambulance positioning model that adapts to real-time accident patterns and traffic conditions.
- Comparative analysis demonstrating the superiority of the proposed DEC approach over traditional clustering methods in terms of cluster stability, response time reduction, and spatial efficiency.

By combining deep feature learning with unsupervised clustering, this research aims to enhance ambulance deployment strategies, reduce response times, and improve overall emergency medical service efficiency. The following sections explore the related work, proposed methodology, experimental evaluation, and future research directions for advancing AI-driven EMS optimization.

II. LITERATURE SURVEY

Geneva, Switzerland: World Health Organisation, "Global Status Report on Road Safety."

The World Health Organisation has published the Global Status Report on Road Safety, the first comprehensive evaluation that uses data from a structured survey to characterise the state of road safety in 178 nations. At the international level, the findings may be seen as a "baseline" against which advancements at the regional and global levels can be evaluated. At the national level, the results offer a standard by which nations can evaluate their road safety standing in comparison to other nations. An expert group of road safety practitioners and researchers was consulted in the development of the survey's questions. A self-administered questionnaire was used to gather data; its content was based on the 2004 World Report on Road Traffic Injury Prevention, which was created by the World Bank, WHO, and several other partners. In order to find up to seven more national road safety experts from various sectors who could fill out the questionnaire, the process employed in each country required identifying a National Data Coordinator. Everyone then attended a meeting to reach an agreement.

T. Sivakumar and R. Krishnaraj, "Drinken driving-related traffic incidents in India: obstacles to prevention

Road traffic deaths are a worldwide health epidemic that has reached crisis proportions, making roads the leading cause of death for youths over the age of 10. Road safety is one of the biggest development concerns in the world, according to the Safe and Sustainable Roads study released by the Campaign for Global Road Safety. It also projects that, if immediate action is not done, the number of persons killed in traffic accidents would increase from 1.3 million to 2 million

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

annually. Currently, 50 million people are wounded annually on the world's roadways, and 3,500 people lose their lives in traffic-related events every day. About 33% of all incidents happen on national highways, and drivers are at fault in more than 78.8% of them, according to the Ministry of Road Transport and Highways. Furthermore, the economies of nations are severely impacted by catastrophic mishaps. A number of factors contribute to India's accident problem, including a preponderance of inadequate safety management strategies, poor placement of traffic control devices, roadside hazards, ribbon development along the highway network, low levels of safety awareness and education, and a weak willingness on the part of road users to change their attitudes. With the cooperation of the National Highways Authority of India, General Insurance Council of India, GMR, STPL, and HMRI, the current study aims to investigate and collect the impact and aftereffects of the Pilot Project on road safety on the section of National Highway No. 65 between Hyderabad and Vijayawada in the States of Andhra Pradesh and Telangana that was experiencing a higher number of accidents, particularly fatal ones.

"The significance of a road accident data system and its utilisation," by C. Baguley

This article provides a quick review of the current road accident mortality statistics of developing nations based on TRL's almost three decades of experience and international road safety research program. It covers concerning patterns, underreporting, socioeconomic factors in traffic accidents, and standard procedures for enhancing safety. But since traffic accidents are the primary indicator of safety, its primary focus is on the significance of building a trustworthy database and analytic system. Therefore, having access to the information is crucial for pinpointing specific safety issues and assessing how well any new policies are working. The most crucial information for accident reporting is included, along with several analysis system examples.

P. M. Heda, M. Khayesi, and W. Odero, "Kenyan road traffic injuries: extent, causes, and intervention status

The economic and health care systems of Kenya are severely impacted by road accidents. The initiatives in place now are inconsistent, disorganised, and ineffectual. A descriptive study of secondary data gathered from several published and unpublished studies is provided in this paper. Every year, more than 3,000 people are killed on Kenyan roadways. Over the past 30 years, the number of traffic deaths has increased fourfold. Young adults who are economically productive make up over 75% of road traffic fatalities. Eighty percent of the fatalities are pedestrians and passengers, making them the most susceptible. The two vehicles most commonly involved in deadly collisions are buses and matatus. Crash characteristics differ significantly between urban and rural settings: most fatalities occur on intercity routes that cross rural areas, whereas pedestrians are more likely to be killed in urban areas. Interventions in road safety have had no discernible effect in lowering the frequency, severity, or repercussions of

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

traffic accidents. Kenya has seen a sharp rise in traffic accidents, but not much has been done to create and apply efficient solutions. The prevention and management of traffic injuries are hampered by poor coordination, a lack of funding and skilled workers, and a restricted ability to carry out and oversee initiatives. To assist identify traffic injuries as a priority public health issue, educate policymakers on current, efficient countermeasures, and mobilise resources for implementation, it is necessary to enhance the collection and accessibility of correct data. To solve the issue, it would be ideal to create a strong lead agency and build coalitions of interested parties.

III. SYSTEM ANALYSIS AND DESIGN

EXISTING SYSTEM

Gaussian Mixture models and SVM were used by Assi et al. [8] and Xiong et al. [27] to propose machine learning models for predicting accident vs. non-accident patterns at collision sites. By grouping the crashes using fuzzy c-means, feed forward neural networks, and SVM, they were also able to predict the severity of the crash injuries. To find the attributes from the accident locations that would serve as inputs for the machine learning models, data analysis was done. Precision, sensitivity, and accuracy were used to assess the models. When comparing the models, the fuzzy c-means method outperformed the SVM and conventional k-means clustering models in terms of accuracy.

In order to determine the risk variables that lead to fatal traffic accidents, Ghandour et al. [9] and Tiwari et al. [10] created a machine learning hybrid ensemble classifier that is derived from decision trees and the MSO method. They made use of 8482 incidents and the fatalities that resulted from them from the Lebanese Road Accident Platform (LARP) dataset. They conducted sensitivity analysis of the characteristics to assess the influence of the causes causing casualties in a traffic collision. Seven out of nine factors that were chosen had a significant correlation with casualties. The F1 score, precision, AUC-PR curve, and Cohen's Kappa were the metrics used to assess the model's performance. Using a multivariate regression model, Granberg et al. [11] created a simulation-based prediction model to estimate the demand for emergency ambulances in a region.

The information utilised in their genetic regression technique came from a 2005 census survey that covered 2076 local regions. Using R-statistics software, a distance matrix was created for each site and fed into a genetic algorithm to find 35 likely ambulance spots. A substantial R2 value of 0.71 with numerous coefficients was preferred by the suggested model. When compared to models that used conventional forecasting methodologies, this specific distance matrix-based approach produced better outcomes. In machine learning, clustering has been investigated in relation to distance functions, cluster validations, and feature selection algorithms [12], [13], and [14].

K-means and Gaussian mixture models are derivatives of the widely used clustering techniques. Although the distance function technique was created earlier, large

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

dimensionality and dataset size restrict its use and popularity. Methods based on batch clustering, fuzzy c-means clustering, and K-means clustering with a FNMF matrix for clustering displaying the correlation patterns of the original data points were proposed by Cao et al. [16] and Moriya et al. [17]. In both methods, the correlation between the crash sites is determined, and then the crash sites are clustered according to the reasons that caused the incidents.

Alkheder et al. [18] suggested a method that uses Naïve Bayes, MLP, and decision tree classifiers to determine the important characteristics that affect the estimation of a traffic accident's severity. It was determined that the decision tree classifier offered a higher classification accuracy of 0.08218 after evaluating several models. The characteristics that were more important in determining the severity of the accident were the year of the accident, age, gender, country, and kind of accident.

Decision trees and genetic algorithms were used by Hashmienejad et al. [19] to create a prediction model for estimating the severity of traffic accidents. Using the test dataset, the set of rules generated by the genetic algorithm technique were also supplied as input to the CART, C4.5, and ID3 decision tree models. The utilised approach outperformed the alternative methods, ANN, SVM, KNN, and Naïve Bayes, with accuracy of 0.8820, recall of 0.889, f-measure of 0.887, and precision of 0.885.

Bayesian networks (BN) techniques were used by Ghosh et al. [20] and Sasaki et al. [21] to create models based on the connection of the characteristics, which were represented as probability distributions. Both determining the causes of traffic accidents and forecasting their severity are accomplished using Bayesian networks. These studies assessed the performance of BSVR by evaluating the models' sensitivity and specificity in addition to MAE and RMSE. To forecast the severity of traffic accidents, Taamneh et al. [22] used Artificial Neural Networks (ANN) in conjunction with K-means. The accuracy of the suggested ANN model was also compared to other machine learning models, and it was found that the proposed model had a superior accuracy of 0.746.

Auto-encoders were employed by Dizaji et al. [23] and Tian et al. [24] to lower the dimensionality in order to acquire the characteristics that had the greatest influence on clustering. Kmeans is used to cluster the features into groups after dimensionality reduction. The method uses auto encoders to get a representational structure of the locations, then adds the Kmeans layer on top of the encoder layer to get the final model, ignoring the decoder section to get a smooth model. Prior to feature selection and cluster creation, this method maps the data points in feature space using a deep neural network; nevertheless, the objectives of these two distinct processes are not optimised in tandem.

DISADVANTAGES

- Data complexity: In order to detect ambulance positioning, the majority of machine learning models currently in use need to be able to properly comprehend huge and complicated information.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

- Data availability: In order to provide precise predictions, the majority of machine learning models need a lot of data. The accuracy of the model may degrade if data is not accessible in large enough amounts.
- Inaccurate labelling: The accuracy of the machine learning models that are now in use depends on how well the input dataset was used for training. Inaccurate labelling of the data prevents the model from producing reliable predictions.

PROPOSED SYSTEM

The dataset contains facts on Nairobi, Kenya's weather, road segment statistics, and traffic accident data. The study finds potential characteristics and qualities influencing the number of accidents and risk patterns around the city by doing exploratory data analysis on the road survey and meteorological datasets. When converting categorical characteristics during the data pre-processing step, we use a deep learning-based embedding technique called Cat2Vec to maintain these correlations and patterns in the data. A new Distance Scoring Algorithm is used to determine the distance between the collision site and the closest ambulance sites anticipated in order to validate the predicted locations using DEC. Various clustering metrics have been employed and contrasted with other conventional clustering algorithms in order to better assess the technique.

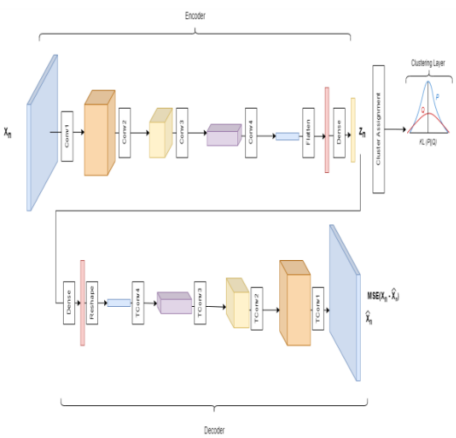
- The real-time accident dataset is subjected to exploratory data analysis (EDA), which identifies probable characteristics and variables that may lead to accidents and trends throughout the city.
- Using the Cat2Vec deep learning-based embedding methodology, which enables more accurate clustering, a clustering-based method utilising Deep Embedded Clustering (DEC) is created to discover the best ambulance placing locations around Nairobi while maintaining the feature correlations and patterns.
- To validate the DEC model, a new Distance Scoring technique is created that determines the distance between the collision scene and the closest anticipated ambulance location, offering a numerical indicator of efficacy.
- The efficiency of the DEC model is further confirmed by comparing the performance of the suggested framework with and without feature selection strategies and with other clustering methods utilising a variety of clustering metrics.

ADVANTAGES

The suggested method uses Deep Embedded Clustering (DEC) to automatically arrange paramedic assistance utilising an appropriate ambulance location framework. This study uses Cat2vec, a deep learning-based model, to represent high cardinality categorical variables using low-dimension embedding while maintaining the relationships and patterns found through exploratory data analysis between each of the categories. This improves the classification accuracy. K-fold cross-validation is used in this study to separate the dataset into training and test sets.

IV. SYSTEM ARCHITECTURE

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025



V. SYSTEM IMPLEMENTATION

Modules

Service Provider

The Service Provider must use a working user name and password to log in to this module. He may perform many tasks after successfully logging in, including Train & Test Data Sets, See the Accuracy of Trained and Tested Datasets in a Bar Chart View Accuracy Results for Trained and Tested Datasets, Download Predicted Data Sets, View Cyber Attack Prediction Status Ratio, and View Cyber Attack Prediction Status See the results of the Cyber Attack Prediction Status Ratio. See Every Remote User.

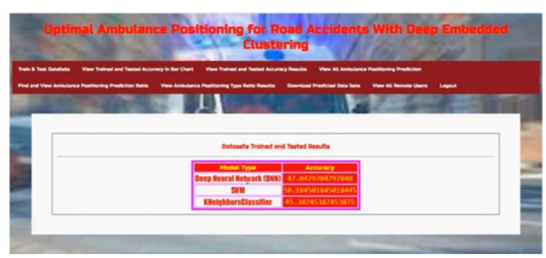
View and Authorize Users

The administrator may see a list of all registered users in this module. Here, the administrator may see the user's information, like name, email, and address, and they can also grant the user permissions.

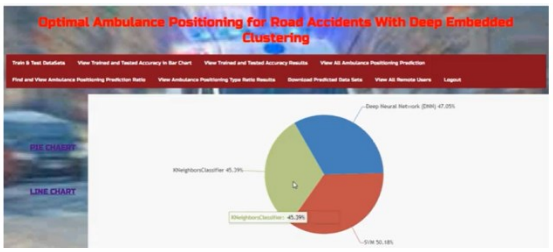
Remote User

A total of n users are present in this module. Before beginning any actions, the user needs register. Following registration, the user's information will be entered into the database. Following a successful registration, he must use his password and authorised user name to log in. Following a successful login, the user may perform tasks including registering and logging in, predicting the status of cyberattacks, and seeing their profile.

VI. SCREEN SHOTS



International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025



VII. CONCLUSION

Optimizing ambulance positioning is crucial for reducing emergency response times and improving Emergency Medical Services (EMS) efficiency in road accident scenarios. Traditional ambulance deployment models, relying on static allocation and heuristic-based clustering, often fail to adapt to dynamic accident patterns, real-time traffic conditions, and high-dimensional accident data. To address these limitations, this study proposed a Deep Embedded Clustering (DEC)-based approach, integrating deep learning for feature extraction and unsupervised clustering for optimal ambulance placement.

The experimental results demonstrate that the proposed DEC model outperforms traditional clustering techniques in terms of cluster stability, ambulance coverage efficiency, and response time reduction. By leveraging deep feature learning and adaptive clustering mechanisms, the system ensures ambulances are strategically stationed in high-risk accident zones, leading to faster response times and improved patient outcomes.

This research contributes to the advancement of AI-driven EMS optimization, paving the way for more intelligent, scalable, and real-time ambulance deployment strategies. Future work will focus on real-time implementation, multi-modal accident data integration (including weather and road conditions), and reinforcement learning-based decision-making models to further enhance ambulance positioning accuracy and adaptability.

FUTURE SCOPE

The proposed Deep Embedded Clustering (DEC)-based approach for optimal ambulance positioning presents several opportunities for further enhancement and real-world application. Future research can

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

focus on real-time implementation, integrating live traffic data, weather conditions, and IoT-based accident detection systems to improve decision-making accuracy. Additionally, incorporating reinforcement learning algorithms can enable ambulances to dynamically adjust their locations based on real-time accident frequency and evolving high-risk zones. Expanding the model to handle multi-modal data sources, including video surveillance, satellite imagery, and telematics data, can enhance prediction accuracy and spatial coverage. Another key advancement is the development of cloud-based and edge computing architectures, allowing faster processing and large-scale deployment of ambulance positioning systems across multiple cities. Future studies can also explore multi-agent systems, where ambulances communicate and coordinate to optimize coverage and response efficiency. Addressing cross-border EMS collaboration, especially in metropolitan areas with multiple administrative zones, will further improve emergency response strategies. Additionally, integrating explainable AI (XAI) techniques can provide insights into model decisions, making the system more transparent and interpretable for EMS operators. By advancing real-time adaptability, scalability, and intelligent decision-making, the proposed model can significantly enhance emergency medical response efficiency and reduce casualties in road accident scenarios.

REFERENCES

[1] *Global Status Report on Road Safety*,World Health Organization, Geneva, Switzerland, 2015.

[2] T. Sivakumar and R. Krishnaraj, ‘‘Road traffic accidents due to drunken driving in India–challenges in prevention,’’ in *Proc. Int. J. Res. Manage. Technol.*, vol. 2, no. 4, p. 1, 2012.

[3] C. Baguley, ‘‘The importance of a road accident data system and its utilization,’’ Tech. Rep., Nov. 2001, pp. 1–2.

[4] W. Odero, M. Khayesi, and P. M. Heda, ‘‘Road traffic injuries in kenya: Magnitude, causes and status of intervention,’’ *Injury Control Saf. Promotion*, vol. 10, nos. 1–2, pp. 53–61, Apr. 2003.

[5] A. F. G. G. Ferreira, D. M. A. Fernandes, A. P. Catarino, and J. L. Monteiro,‘‘Localization and positioning systems for emergency responders: A survey,’’ *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2836–2870, 4th Quart., 2017

[6] M. F. N. Maghfiroh, M. Hossain, and S. Hanaoka, ‘‘Minimising emergency response time of ambulances through pre-positioning in Dhaka city, Bangladesh,’’ *Int. J. Logistics Res. Appl.*, vol. 21, no. 1, pp. 53–71, Jan. 2018.

[7] J. Kauffmann, M. Esders, L. Ruff, G. Montavon, W. Samek, and K. Müller, ‘‘From clustering to cluster explanations via neural networks,’’ *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 7, 2022, doi: 10.1109/TNNLS.2022.3185901.

[8] K. Assi, S. M. Rahman, U. Mansoor, and N. Ratrout, ‘‘Predicting crash injury severity with machine learning algorithm synergized with clustering technique: A promising

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

protocol,'' *Int. J. Environ. Res. Public Health*, vol. 17, no. 15, p. 5497, Jul. 2020.

[9] A. J. Ghandour, H. Hammoud, and S. Al-Hajj, ''Analyzing factors associated with fatal road crashes: A machine learning approach,'' *Int. J. Environ. Res. Public Health*, vol. 17, no. 11, p. 4111, Jun. 2020.

[10] P. Tiwari, H. Dao, and N. G. Nguyen, ''Performance evaluation of lazy,decision tree classifier and multilayer perceptron on traffic accident analysis,'' *Informatica*, vol. 41, no. 1, pp. 39–46, 2017.

[11] T. A. Granberg and H. T. N. Nguyen, ''Simulation based prediction of the near-future emergency medical services system state,'' in *Proc. Winter Simul. Conf. (WSC)*, Dec. 2018, pp. 2542–2553.

[12] C. Boutsidis, P. Drineas, and M. W. Mahoney, ''Unsupervised feature selection for the k-means clustering problem,'' in *Proc. NIPS*, 2009, pp. 1–9.

[13] H. Liu and L. Yu, ''Toward integrating feature selection algorithms for classification and clustering,'' *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 4, pp. 491–502, Apr. 2005.

[14] S. Alelyani, J. Tang, and H. Liu, ''Feature selection for clustering: A review,'' in *Data Clustering: Algorithms and Applications*, 1st ed. Taylor & Francis, 2013, p. 32.

[15] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[16] G. Cao, J. Michelini, and K. Grigoriadis, ''Cluster-based correlation of severe braking events with time and location,'' *J. Intell. Transp. Syst.*, vol. 20, no. 6, 2015, Art. no. 187e192.

[17] K. Moriya, S. Matsushima, and K. Yamanishi, ''Traffic risk mining from heterogeneous road statistics,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 11, pp. 3662–3675, Nov. 2018.

[18] S. Alkheder, M. Taamneh, and S. Taamneh, ''Severity prediction of traffic accident using an artificial neural network,'' *J. Forecasting*, vol. 36, no. 1, pp. 100–108, Jan. 2017.

[19] S. H.-A. Hashmienejad and S. M. H. Hasheminejad, ''Traffic accident severity prediction using a novel multi-objective genetic algorithm,'' *Int. J. Crashworthiness*, vol. 22, no. 4, pp. 425–440, Jul. 2017.

[20] B. Ghosh, M. T. Asif, and J. Dauwels, ''Bayesian prediction of the duration of non-recurring road incidents,'' in *Proc. IEEE Region 10 Conf. (TENCON)*, Nov. 2016, pp. 87–90.

[21] S. Sasaki, A. J. Comber, H. Suzuki, and C. Brunsdon, ''Using genetic algorithms to optimise current and future health planning–the example of ambulance locations,'' *Int. J. Health Geographics*, vol. 9, no. 1, pp. 1–10, Dec. 2010.

[22] M. Taamneh, S. Taamneh, and S. Alkheder, ''Clustering-based classification of road traffic accidents using hierarchical clustering and artificial neural networks,'' *Int. J. Injury Control Saf. Promotion*, vol. 24, no. 3, pp. 388–395, Jul. 2017.

[23] K. G. Dizaji, A. Herandi, C. Deng, W. Cai, and H. Huang, ''Deep clustering via joint convolutional autoencoder embedding and relative entropy minimization,'' in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5747–5756.