

Deep Neural Network Based Complex-Heterogeneous Cyberspace Cartographic Visualization

Lan Zhang¹, Zhiyong Liu², Guangxia Wang¹, Lei Xiong^{2*}

Abstract

It seems complicated to imagine our life without the Internet, but it is only a few decades old, and the Internet forms the most extensive cyberspace environment. Cyberspace is the non-physical environment created by joint computers inter-operating on a network. The complex heterogeneous cyberspace is a virtual computer world that encompasses a variety of computers, network devices, and systems that have been manufactured by different entities. As an essential spatial cognitive tool, the map has significantly contributed to human civilization for thousands of years. However, the cartographic elements of traditional maps are mostly geospatial entities or phenomena, and only some people apply them to draw abstract and virtual cyberspace resources, resulting in the development of cyberspace cartography lagging. Additionally, the process of data visualization involves the conversion of data into visual forms such as charts and graphs, with the aim of effectively conveying the data's importance. The cartographic visualization is to realize the visual expression of cyberspace, which is an essential basis for understanding cyberspace. Therefore, this paper introduces a deep neural network (DNN) to study complex-heterogeneous cyberspace cartographic visualization. At first, locally linear embedding is adopted to reduce the data dimensionality. Then, DNN is used to train the cartographic coordinates data obtained after data dimensionality reduction. Finally, several different data modeling methods are studied concerning temporal and spatial attributes to achieve complex-heterogeneous cyberspace cartographic visualization effectively. The simulation is powered by datasets sourced from cyberspace, which are accessible via data.world. The results of the simulation demonstrate that the suggested approach is superior in efficiency when compared to the baseline methods.

Key Words: Cyberspace, Complex-Heterogeneous, Cartographic Visualization, DNN, LLE.

I. INTRODUCTION

Today online world is our new home, many of our activities, in fact most of our activities like many students are going online classes, many of us actually book tickets online, most of our people actually also doing online shopping, and the communities are also conducting online conferences. In recent decades, networks have expanded from being a simple means of communication to being an integral part of nearly every aspect of our modern lives. The scale, speed, and dynamic complexity of networks have all increased. Cyberspace is what the Internet has become due to the widespread usage of computers and networking technology. Cyberspace, or the virtual environment enabled by computers and networks, is an environment where people all over the world may talk to each other using the Internet and other electronic means [1-2]. As a new frontier of people, cyberspace has reached a broad consensus. How to bet-

ter explore and represent cyberspace has attracted wide attention around the world. Since 2012, the United States had launched the "Treasure Map Project" and the "mission X" to study real-time and interactive global Internet maps, which provided basic support for research on cyberspace information [3]. In 2014, Kaspersky lab in Russia released its new interactive cyber threats real-time map, which was dedicated to real-time representation of network activities [4]. In 2018, China Internet Security Conference (CISC) showed the smart city system, which integrated basic cyber information and showed it in a centralized way, so as to monitor and respond to cyber security attacks in real time [5].

Given the heterogeneous network paradigm [6] and different applications such as terrestrial communication network [7], mobile marine network [8], aerostat platform [9] and satellite communication networks [10], the form of cyberspace is often complex and heterogeneous, so the re-

Manuscript received April 27, 2023; Revised June 02, 2023; Accepted June 02, 2023. (ID No. JMIS-23M-04-016)

Corresponding Author (*): Lei Xiong, +86-18971223361, 2011302590228@whu.edu.cn

¹Institute of Geospatial Information, Information Engineering University, Zhengzhou, China, ranhou0611@163.com, wangguangxia2011@163.com

²School of Remote Sensing and Information Engineering, Wuhan University, Wuha, China, bear2001321@126.com, 2011302590228@whu.edu.cn

search of this paper is based on complex-heterogeneous cyberspace [11].

Visualization is a graphical representation that uses appropriate graphs, charts and maps to represent the relationships between data [12-13]. As a technical method of spatial cognition and representation, cartographic visualization is not only limited to the surface space, but also can be extended to the space beyond the surface [14-15]. Visualization research is very important, and ground-based Internet cartographic visualization is even more important. While geographical data is related to spatial location and closely related to visualization. Cyberspace cartographic visualization can be applied to typical business scenarios such as real-time monitoring of cyberattacks, traceback of cyber security events, communication support, and so on. So it is obviously significant for the research on complex-heterogeneous cyberspace cartographic visualization.

In traditional geographical space, map as an important carrier of describing geographical phenomenon, which has been an indispensable tool for operational command since ancient times [16]. The traditional method of data visualization analysis is to draw corresponding graphs to make people better understand the data. Currently, it is needed for cyberspace map that can fully show the information of cyberspace, so as to establish the connection between cyberspace and geographical space. The ultimate goal of the cyberspace visualization is to fully show cyber information in the form of cyberspace map and realize the visualization and digitization of cyberspace, and thus provide intuitive and valuable information for decision-makers to reduce the uncertainty of decision-making. The early stage of the development of cyberspace visualization is faced with problems such as weak theoretical basis and immature technology, while the introduction of big data and artificial intelligence (AI) provides a new perspective for the research of cyberspace visualization [17].

Although computer vision and multimedia systems are not directly related to cartographic visualization, there are some applications of computer vision techniques that can be useful for processing and analyzing visual data in cartographic visualizations. For example, computer vision techniques such as object detection and tracking can automatically identify and label features on a map, such as roads, buildings, and bodies of water. Image segmentation techniques can be used to separate different regions or layers of a map, such as land and water or different types of land use. Machine learning techniques can also be applied to classify and predict patterns in geographic data. For example, clustering algorithms can group similar features on a map, such as neighborhoods or land use types. Classification algorithms can predict the likelihood of certain features or events occurring in a particular location, such as the likelihood of flooding in a particular area. Additionally, multi-

media systems can play a role in creating interactive and immersive cartographic visualizations that allow users to explore and interact with geographic data in new ways. For example, virtual reality and augmented reality technologies can create immersive experiences allowing users to explore geographic data in three dimensions. While cartographic visualization may not be directly related to computer vision or multimedia systems, some applications of these techniques can be helpful in processing and analyzing visual data in cartographic visualizations and creating interactive and immersive experiences for users.

Deep learning is a form of machine learning that takes its cues from the structure of the human brain; in the context of deep learning, this structure is known as an artificial neural network [18]. Deep learning has the advantages such as well learning ability and data-driven with high threshold. While along with various advantages of neural networks, the most common ones are classification and cluster. The strong combination of deep learning and neural networks is deep neural network (DNN), which happens to be close to the problem of cyberspace visualization.

LLE is a dimensionality reduction technique that focuses on capturing the local structure of the data. It seeks a lower-dimensional data representation to preserve the pairwise distances between neighboring points. The preservation of local relationships allows for a meaningful visualization of the data while reducing the dimensionality. LLE can be particularly valuable in complex and heterogeneous cyber data. Cyber data often consists of high-dimensional and diverse attributes, such as network connections, protocols, time-stamps, and other relevant features. Applying LLE can effectively reduce the dimensionality of this data while retaining its essential characteristics, enabling a more concise and meaningful representation that captures the underlying structure and relationships of the data. First, it helps overcome the curse of dimensionality by reducing the dimensionality of the data, which can be beneficial for visualization purposes. Second, LLE preserves the local structure of the data, allowing for a more accurate representation of the data's inherent relationships. It is imperative in complex and heterogeneous cyber data, where understanding the local interactions and dependencies is crucial. Integrating vector autoregressive moving average (ARMA) for spatiotemporal data modeling is another unique aspect of their approach. ARMA models are widely used for time series analysis, and by applying it to the spatiotemporal aspects of the data, it demonstrates their ability to capture and analyze temporal patterns and dependencies within the network. This integration enables a comprehensive understanding of how the network evolves, adding an essential temporal component to their visualization technique. Furthermore, the use of DNN for training is noteworthy. DNN has

demonstrated exceptional capabilities in learning complex representations and patterns from data. By employing a DNN in the complex heterogeneous cyber cartographic visualization, we leverage the power of deep learning to extract meaningful features and representations from the cyber data, enabling a richer and more accurate visualization of the network.

The paper is structured as follows. Section 2 reviews the related work. In Section 3, we study the DNN based complex-heterogeneous cyberspace cartographic visualization. The experimental results are shown in Section 4. Section 5 concludes this paper.

II. RELATED WORK

In today's digital age, we are continually bombarded with information, most of it may or may not be reliable. Although raw data is used to determine if something is genuine or incorrect, it is rarely presented to the public. It is easy to understand how rows upon rows of numbers may be difficult to interpret. Because of this, we commonly use data visualization to present patterns and trends in data more easily. In [19], the authors presented a novel distributed union-find algorithm that features asynchronous parallelism and k-d tree-based load balancing for scalable visualization and analysis of scientific data. In [20], the authors proposed a new perspective of ensemble data analysis using the attribute variable dimension as the primary analysis dimension. Using matplotlib callbacks, visualization toolkits, and embedded HTML visualizations, the authors of [21] demonstrated three methods for incorporating interactive visualizations into Jupyter Notebooks. In [22], the authors introduced the interactive catchment explorer, a web-based interactive data visualization platform for investigating environmental information and model results. For the purpose of visualizing abstract gaze data, the authors of [23] presented a data processing approach based on gaze behavior. To accurately incorporate information about transcriptomic variability into the visual interpretation of single-cell RNA sequencing data, the authors of [24] presented den-stochastic neighbor embedding and dens manifold approximation and projection, density-preserving visualization tools based on t-stochastic neighbor embedding and uniform manifold approximation and projection, respectively. A novel graphical tool for the visualization of health data was published in [25], which may be used to quickly monitor patients' health condition remotely. In [26], the authors suggested a new supervised dimension-reduction approach termed supervised t-distributed stochastic neighbor embedding, which achieved dimension reduction while maintaining the similarities between data points in both the feature and outcome spaces. The suggested technique can handle high-di-

mensional data, making it useful for both prediction and visualization applications. Combining illustration with data visualization was investigated in [27], where the authors presented interactive picture segmentation and gridding techniques. While many studies have focused on visualizing data in a homogeneous network, complicated heterogeneous cyberspace has received far less attention.

The term "cyberspace visualization" refers to the use of visual language to explain and analyze a wide range of cyberspace phenomena and occurrences, such as the visualization of network elements, network structure, and security incidents. Cyberspace information system was described in [28] as a parallel to geographic information systems, with the latter allowing for visualization based on a geographical coordinate system. This led to the proposal of a multi-dimensional and multi-view cyberspace information system model. In [29], the authors created an architecture for visualizing the cyber battleground from border gateway protocol archive data, which included border gateway protocol connection information data from routers all over the world. In [30], the capability demand of the joint operation for cyberspace war scenario visualization system was used to do the system function analysis. In [31], two generalization approaches were presented after analyzing and measuring different forms of characteristic information of point cluster characteristics in cyberspace from four perspectives: statistics, metrics, topology, and themes. In [32], an ontology-based knowledge representation method for cyberspace situational information elements was proposed; this study aided in the understanding, modelling, and presentation of the cyberspace environment, and it served as a useful point of reference for the study of related technologies. In [33], the authors suggested the connotation and technological route of cyberspace visualization based on the idea of the "man-land-network" nexus and explain the visualization of cyberspace elements, cyberspace relations, and cybersecurity incidents. In [34], considering that the distance cartogram could express the characteristics of spatial relational information in a simplified and deformed geographic space, a composite distance cartogram was designed according to the cyberspace information visualization model. To generate a cyberspace composite distance cartogram, the coordinate transformation principle and method for the nodes of network communities were proposed. In [35], from the basic concept of cyberspace, based on the geospatial information grid, the authors studied the network space physical domain, logic domain and social domain partition method respectively. In [36], according to the spatial correlation degree of cyberspace and its elements, mapping methods for cyberspace were classified and the key technologies that needed to be solved were proposed. To the best of our knowledge, there are almost no researches on

using DNN to realize cyberspace visualization, even with regard to complex-heterogeneous cyberspace. However, the research on DNN based complex-heterogeneous cyberspace is an essential part in cartographic visualization, which motivates this paper.

III. CYBERSPACE CARTOGRAPHIC VISUALIZATION

In light of the complexity and heterogeneity of cartographic data in complex-heterogeneous cyberspace, it is necessary to reduce the data dimensionality and greatly reduce the workload of DNN. Dimensionality reduction process reduces the number of random variables or features under consideration in a machine learning algorithm. Data dimensionality reduction refers to the process of mapping a sample from a high-dimensional space to a low-dimensional space through linear or nonlinear mapping to obtain a meaningful low-dimensional representation of high-dimensional data.

3.1. Data Dimensionality Reduction

Data dimensionality reduction methods are commonly divided into linear methods and nonlinear methods, which can be used for data visualization. Although the linear method is simple to calculate, it cannot find the nonlinear regularity of distribution of data, especially for the manifold distribution data. In order to overcome the problem of nonlinear distribution of data in this paper, LLE is introduced to reduce the data dimensionality. LLE unfolds the nonlinear manifold in a piece-wise manner. Each piece is unfolded and the unfolded pieces are put together to have the entire unfolded manifold. The steps of LLE is summarized as follows.

- (1) Select neighbors and construct k -nearest neighbors (k NN) graph.
- (2) Reconstruct with linear weights. Find the reconstruction weights for each point based on their neighbors.
- (3) Map to embedded coordinates. Use the obtained weights to embed the points in the low dimensional subspace.

The T -dimensional training dataset is represented by $X = (x_1, x_2, \dots, x_q) \in \mathbb{R}^{T \times q}$, while t -dimensional training dataset after dimensionality reduction is represented by $Y = (y_1, y_2, \dots, y_q) \in \mathbb{R}^{t \times q}$. The distance between x_m and x_n is denoted as $dis(x_m, x_n)$.

A k NN graph is formed using pairwise Euclidean distance between the data points. Therefore, every data point has k neighbors. Let P_m denote the adjacent points set. Then compute the linear reconstruction coefficient ω_m for x_m .

$$\begin{aligned} \min_{\omega_m} \sum_{m=1}^q ||x_m - \sum_{n \in P_m} \omega_{mn} x_n||^2. \\ \text{s.t. } \sum_{n \in P_m} \omega_{mn} = 1, m = 1, 2, \dots, q. \end{aligned} \quad (1)$$

The linear reconstruction coefficient ω_m is kept unchanged, and the lower dimensional space coordinate y_m corresponding to x_m is solved. The constraint is $\sum_{m=1}^q y_m = 0$ and $\sum_{m=1}^q y_m y_m^T = qI$, where I is the identity matrix.

$$\min_{y_m} \sum_{m=1}^q ||y_m - \sum_{n \in P_m} \omega_{mn}||^2. \quad (2)$$

The core of LLE algorithm is the establishment of reconstruction coefficient. Let $\Phi = [x_{P_m(1)}, x_{P_m(2)}, \dots, x_{P_m(k)}] \in \mathbb{R}^{T \times k}$ denote the matrix composed of k nearest neighbors of x_m , and the matrix composed of k x_m is represented by $X = (x_1, x_2, \dots, x_q) \in \mathbb{R}^{T \times q}$.

$$||x_m - \sum_{n \in P_m} \omega_{mn}||^2 = ||x_m - \Phi \omega_m||^2 = \omega_m^T M \omega_m, \quad (3)$$

where $M = (X - \Phi)^T (X - \Phi)$, and because $\sum_{n \in P_m} \omega_{mn} = 1$, $x_m = X \omega_m$. Therefore, ω_m can be calculated as follows.

$$\begin{aligned} \min_{\omega_m} \omega_m^T M \omega_m \\ \text{s.t. } \omega_m^T E = 1, \end{aligned} \quad (4)$$

where $E = (1, 1, \dots, 1)^T$, then Lagrange function is constructed as follows.

$$Lag(\omega_m, \lambda) = \omega_m^T M \omega_m - \lambda(\omega_m^T E - 1), \quad (5)$$

where λ is the Lagrange multiplier. Taking the derivative of equation (5) with respect to ω_m and λ , and we have

$$\begin{cases} \frac{\partial Lag}{\partial \omega_m} = 2M\omega_m - \lambda E = 0 \\ \frac{\partial Lag}{\partial \lambda} = \omega_m^T E - 1 = 0 \end{cases}. \quad (6)$$

According to equation (6), and we have

$$\begin{cases} \omega_m = \frac{\lambda}{2} M^{-1} E \\ \lambda = \frac{2}{E^T M^{-1} E} \end{cases}. \quad (7)$$

The calculation of low dimensional coordinates in equation (2) can be embedded by coefficients construction in equation (7).

Given the above, the reconstruction coefficient ω_m of each x_m contains local information in high dimensional data, and the characteristics of the related data can be well preserved in low dimensional data. The value of ω_m determines whether dimensionality reduction can be completed.

3.2. Data Training

DNN is used to train the cartographic coordinates data obtained after data dimensionality reduction to reduce the workload of DNN.

During DNN training, a certain number of samples are used to form a Mini-batch. Supposing that the dimension of each sample is T and the number of samples is Q , then each Mini-batch is a matrix of $T \times Q$. The whole training process will go through three parts: forward computation, backward computation and weight updating. Assuming that the layer number of neural network is L . In forward computation process, the output layer O after training is obtained through input Mini-batch computation, and the process can be expressed as follows.

$$\begin{cases} y_{l,n}^{(1)} = \text{relu}(\sum_{m=1}^T x_{l,m} \times w_{n,m}^{(1)} + b_n^{(1)}) \\ y_{l,n}^{(\gamma)} = \text{relu}(\sum_{m=1}^{Q/T-1} y_{l,m}^{(\gamma-1)} \times w_{n,m}^{(\gamma)} + b_n^{(\gamma)}) \\ z_{l,n} = \text{softmax}(\sum_{m=1}^{Q/T-1} y_{l,m}^{(\gamma-1)} \times w_{n,m}^{(\gamma)} + b_n^{(\gamma)}) \end{cases}, \quad (8)$$

where x is the input Mini-batch data, that is, the data of input layer. y^γ is the hidden layer result of γ th layer. $w^{(\gamma)}$ and $b^{(\gamma)}$ are the weight and bias of γ th layer network respectively, and z is the output value of forward computation. Additionally, $\text{relu}()$ and $\text{softmax}()$ are forward activation functions corresponding to two different operations in forward computation process respectively.

3.3. Data Model Design

Data of cartographic in complex-heterogeneous cyberspace mainly includes coordinates and other data such as device, application, data, IP, protocol and subject of network. Several data different modeling methods are described in terms of temporal and spatial attributes in order to realize complex-heterogeneous cyberspace cartographic visualization.

3.3.1. Time Series Data Modeling

There are many time series metrics in multidimensional data. Time series is basically a sequence where we record a metric over regular intervals. For a certain kind of network security event in complex-heterogeneous cyberspace, the risk distribution of such event can be forecasted by these metrics with time attribute. The following models are mainly used, which model to choose in practical application depends on the time fluctuation and dependence of data.

3.3.1.1. ARMA model

The ARMA model, which combines the AR model and the MA model, is a crucial tool for studying time series. Forecast index data across time is treated as a random sequence by ARMA. This collection of random variables is

dependent on one another, which symbolizes the timelessness of the original data. Assuming that the impact factors are u_1, u_2, \dots, u_k , and the forecast object can be obtained by regression analysis.

$$Y_t = b_1 u_1 + b_2 u_2 + \dots + b_p u_p + E_t, \quad (9)$$

where Y is the observed value of the forecast object. E_t is the error. As a forecast object, and Y_t is affected by its own changes according to the following equation.

$$Y_t = b_1 Y_{t-1} + b_2 Y_{t-2} + \dots + b_p Y_{t-p} + E_t. \quad (10)$$

Error has dependencies in different phases, which can be defined as follows.

$$E_t = \varepsilon_t + \mu_1 \varepsilon_{t-1} + \mu_2 \varepsilon_{t-2} + \dots + \mu_q \varepsilon_{t-q}. \quad (11)$$

Thus, the expression of ARMA model can be defined as follows.

$$Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \dots + b_p Y_{t-p} + \varepsilon_t + \mu_1 \varepsilon_{t-1} + \mu_2 \varepsilon_{t-2} + \dots + \mu_q \varepsilon_{t-q}. \quad (12)$$

3.3.1.2. ARIMA model

Auto regressive integrated moving average (ARIMA) model is a forecasting algorithm that takes into account previous past values to forecast future values because it considers that the information is found in those past values can be indicative of future values. The ARIMA model is defined by the three parameters p , d , and q . For a stationary time series, p is the order of the auto regressive term, d is the order of differences, and q is the order of the moving average term. The number of lags in the dependent variable (p), the number of differenced iterations (d), and the number of lags in the error term (q) are all displayed in the ARIMA (p, d, q) model. For instance, the ARIMA model with parameters (1,1,2) contains a one-lag dependent variable (1), a first-difference stationary (1) variable, and a two-lag error term (2). In this case, the ARIMA model comprises a one-lag dependent variable (1), a zero-lag independent variable (0), and a one-lag error term (1). ARIMA (1,0,1) equals ARMA (1,1) if the series is level.

The distinguishing factor between the ARIMA and ARMA models lies in the former's ability to convert non-stationary time series into stationary time series via differential operation, thereby facilitating modelling. The specific equation of ARIMA is the same as ARMA except that a difference operation is added before modeling.

3.3.1.3. ARCH model

Autoregressive conditional heteroskedasticity (ARCH)

model conveys that the series in question has a time-varying variance (*heteroskedasticity*) that depends on (*conditional on*) lagged effects (*autocorrelation*). ARCH model takes all available information as condition and uses the form of auto regression to describe variance variation. For time series, available information is different at different time, and the corresponding conditional variance is also different. ARCH model can be used to describe the conditional variance with time variation.

The basic idea of ARCH model is that under the previous information set, and the occurrence of a noise at a certain time is Gaussian distribution. The mean of the Gaussian distribution is zero and the variance is a quantity that changes over time (i.e., conditional heteroscedasticity). While the time-varying variance is a linear combination of the squares of the past finite term noise values (i.e., auto regression), which constitutes the ARCH model.

Let the error variance be time-varying, that is, heteroscedastic and call it h_t , then the basic ARCH(1) process is defined as follows.

$$h_t = b_0 + b_1 u_{t-1}^2. \quad (13)$$

When a big shock occurs in the previous period $t - 1$, it is more likely that the value of u_t in absolute terms will also be bigger, that is, when u_{t-1}^2 is small or large, the variance of the next innovation u_t will also be small or large. The ARCH(q) model can be defined as follows.

$$h_t = b_0 + \sum_{i=1}^q b_i u_{t-i}^2. \quad (14)$$

The ARCH(q) model shows that the variance or volatility in a given period depends on the magnitudes of the squared errors in the past q periods.

3.3.1.4. GARCH model

To resolve the problem of negative estimates, the generalized ARCH (GARCH) model is developed, which includes the lagged conditional variance terms as autoregressive terms and uses few parameters to capture long lagged effects.

Based on ARCH(q) model, the GARCH (p, q) model can be defined as follows.

$$h_t = b_0 + \sum_{k=1}^p \theta_k h_{t-k} + \sum_{i=1}^q b_i u_{t-i}^2. \quad (15)$$

The GARCH (1,1) model contains one lagged term of the conditional variance (h) and one lagged term of the squared error (u^2).

3.3.2. Spatiotemporal Data Modeling

Each object, event or phenomenon of complex-heterogeneous cyberspace is associated with time and space, result-

ing in a wide range of spatiotemporal application fields. Spatiotemporal data model is the basis of spatiotemporal data management, and the effective processing of spatiotemporal data needs to be based on spatiotemporal database model.

The metric with both spatial and temporal attributes in complex-heterogeneous cyberspace is the coordinates of devices or subject of network. In complex-heterogeneous cyberspace cartographic visualization, the risk distribution of network security event can be forecasted in order to realize the “battle on map”. Taking coordinates in multidimensional data as an example, the change of the coordinates of a node with time is a time series analysis problem, but if the coordinates of other nodes change, the coordinates of the node may also change accordingly. Therefore, the analysis of spatial data is also designed. The specific method is defined as follows.

$$\sum_{k=1}^p \sum_{i=1}^k \Phi_{chc} C^{chc} u_{t-1} = b_t + \sum_{j=1}^q \sum_{i=1}^k \theta_{chc} C^{chc} b_{t-1}, \quad (16)$$

where Φ_{chc} is the autoregressive coefficient of complex-heterogeneous cyberspace, C^{chc} is the coordinate weight matrix, u_{t-1} is the autoregressive term, b_t is the moving regression current deviation, θ_{chc} is the moving regression coefficient, and b_{t-1} is the moving regression deviation term. The model used in this paper is vector ARMA (VARMA) model with spatial dependence.

IV. SIMULATION RESULTS

4.1. Setup

The proposed DNN based complex-heterogeneous cyberspace cartographic visualization is implemented via three parts. First, LLE method is introduced to reduce the data dimensionality. Then, DNN is used to train the data after dimensionality reduction. Finally, in order to realize complex-heterogeneous cyberspace cartographic visualization, the data model is designed with respect to temporal and spatial. The computational simulation is being executed on a computing equipped with an Intel i9-11900k processor operating at a frequency of 3.5 GHz, and a memory capacity of 16 GB with a clock speed of 2,666 MHz. The simulation is driven based on cyberspace datasets available on data.world. In the simulation, three algorithms are selected with comparison: a differential privacy enabled DNN learning framework (DNN-DP) [37], channel state information (CSI)-based DNN (CSI-DNN) [38] and convolutional autoencoder with residual blocks-DNN (CAERES-DNN) [39]. The evaluation of simulation results often involves the utilization of three classification metrics, namely precision ratio (P), recall ratio (R), and F1 score (F). These metrics are commonly employed to compare and assess the

quality of the aforementioned results. The value of P denotes the level of accuracy in the recognition of coordinates. The value of R represents the extent to which the visualization outcomes encompass the cyberspace. The metric F provides a comprehensive assessment of both precision (P) and recall (R).

The model of DNN based complex-heterogeneous cyberspace in simulation has a total of nine layers, among which the number of nodes in the hidden layer is 2048, the number of nodes in the top layer is 8992, and the number of nodes in the input layer is 400. The network parameters are initialized with Gaussian distribution with mean value of 0 and variance of 1. The size of Mini-batch is set to 200.

4.2. Performance Analysis

4.2.1. Data Model Test with DNN Based

Fig. 1 describes the trends of the precision ratio of the model with four algorithms in complex-heterogeneous cyberspace cartographic visualization. Increasing the number of nodes in cyberspace has a negative effect on the accuracy of all algorithms, as shown in Fig. 1. However, the accuracy of the method suggested in this study remains consistently greater than the accuracy of the other three baselines. Compared with sigmoid activation function, the method proposed in this paper with ReLU activation function has three main changes, which are unilateral inhibition, relatively wide excitatory boundary and sparse activation. The compared three baselines with sigmoid activation function have the fatal error that if the initial number of nodes in cyberspace is large, most neurons may be in the saturation state and kill gradient, which will make the network difficult to learn. As a result, the precision ratio of the method suggested in this study is greater than that of the other three baselines, and its theoretical accuracy is higher as well.

In complex-heterogeneous cyberspace, Fig. 2 depicts the variation of recall ratio of four algorithms. As shown in Fig. 2, the recall of the method presented in this study is greater

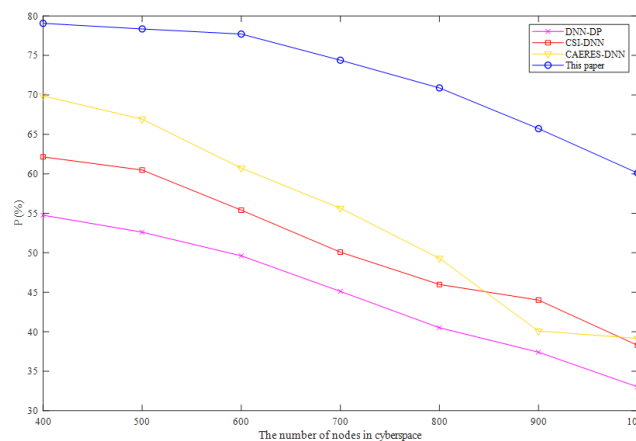


Fig. 1. The comparison of P of the four DNN-based algorithms.

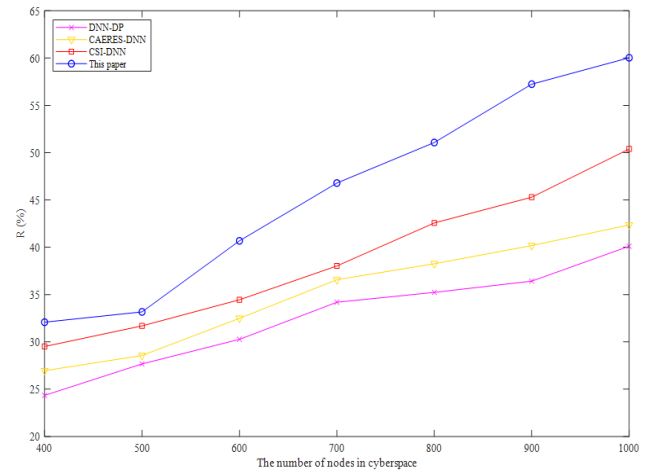


Fig. 2. The comparison of R of the four DNN-based algorithms.

than that of the other three baselines, and it rises as the number of cybernetic nodes rises. In a fully connected DNN structure, connections may be made between the bottom neurons and all of the higher neurons of the other three baselines, increasing the number of parameters and lowering the recall. After dimensionality reduction, the benefits of the method suggested in this study become more apparent when applied to cartographic data in complex-heterogeneous cyberspace.

The F1 score of four algorithms for visualizing complicated heterogeneous cyberspace is shown in Fig. 3. Precision and recall should be as high as feasible when assessing the simulation results. However, in most situations, the two ratios contradict one another. Therefore, the two ratios should be taken into account using the F1 score, which can provide an overall indication of the method's efficacy. I find that the F1 score is greatest for the approach described in this work, followed by DNN-DP, CSI-DNN, and CAERES-DNN. Despite the CSI-DNN's high F1 score and poor precision ratio, keeping features invariant to the investigated impairments is the algorithm's stated goal.

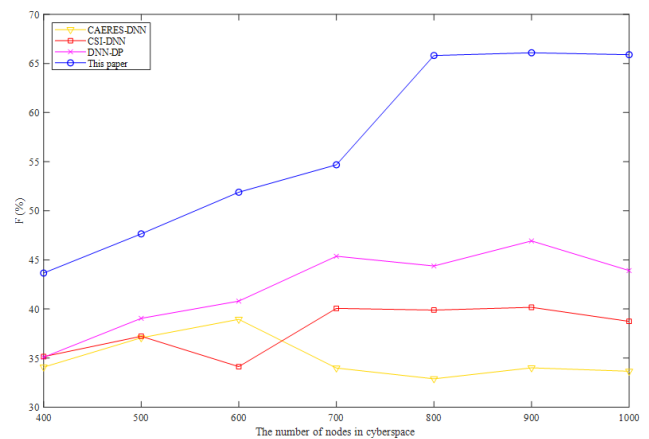


Fig. 3. The comparison of $F1$ of the four DNN-based algorithms.

4.2.2. Data Model Test with Neural Network

In order to verify the proposed method in this paper has a good performance in DNN based complex-heterogeneous cyberspace cartographic visualization. The other neural network algorithms are selected for comparison, which are convolutional auto-encoders and hinge loss CNN (CAE-HL-CNN) [40] and dual path CNN-recurrent neural network (RNN) cascade network (DPCRCN) [41].

As can be seen from Fig. 4, Fig. 5, and Fig. 6, the precision ratio, recall ratio and F1 score of the proposed method are the highest comparing with other three neural network algorithms. CNN and RNN are essentially different though they can conduct sequence modeling. RNN has an order in time dimension, and the order of input will affect the output. CNN mainly obtains the overall information from local information aggregation and extracts the hierarchical information from the input layer. The convolution kernel of CNN emphasizes the window in space, which is similar to the time series problem, but RNN does not consider the spatial cases. The proposed method firstly reduces the dimensionality of high-dimensional data to avoid the pressure of processing high-dimensional data. Although DNN cannot

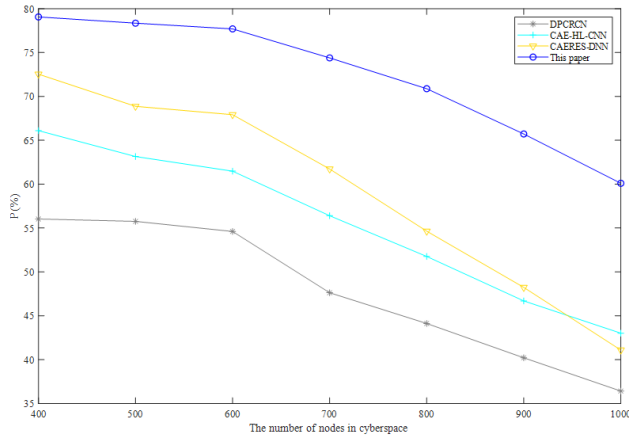


Fig. 4. The comparison of P of the four neural network algorithms.

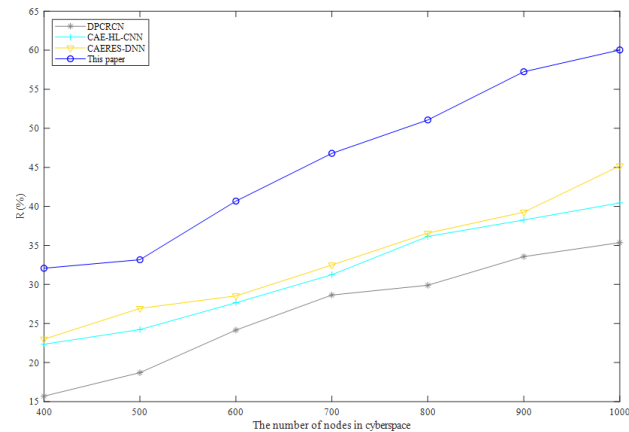


Fig. 5. The comparison of R of the four neural network algorithms.

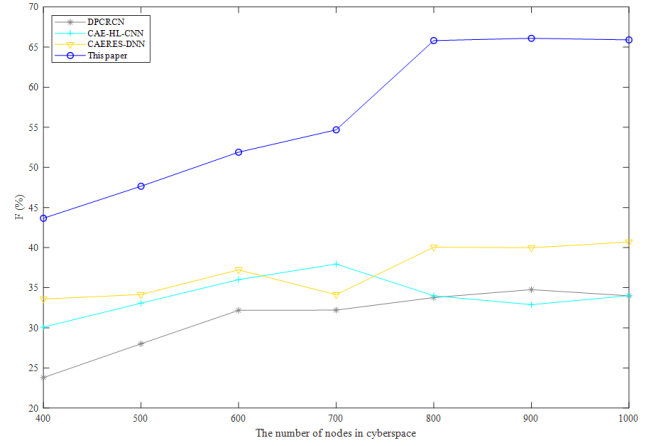


Fig. 6. The comparison of F1 of the four neural network algorithms.

process the changes in time series, the proposed method based on spatiotemporal data modeling which can effectively reduce the impact of the changes in time series.

4.2.3. Training Time

Four DNN-based algorithms are used in training time comparison, and the training time comparison of this paper, CAERES-DNN, CSI-DNN and DNN-DP is reported in Fig. 7. With the increasing number of nodes in complex-heterogeneous cyberspace, the training time for forecasting the risk distribution of network security event is also growing. Even with the growing number of nodes in cyberspace, the training time is surprisingly close to two minutes. With 800 nodes, the training time only increases by a little range, demonstrating the method's superior convergence performance. ReLU activation function is used in this paper, and momentum optimization is added to make the model jump out of locally optimal easily, which is also reduce oscillation and accelerate convergence speed.

4.2.4. Mini-Batch

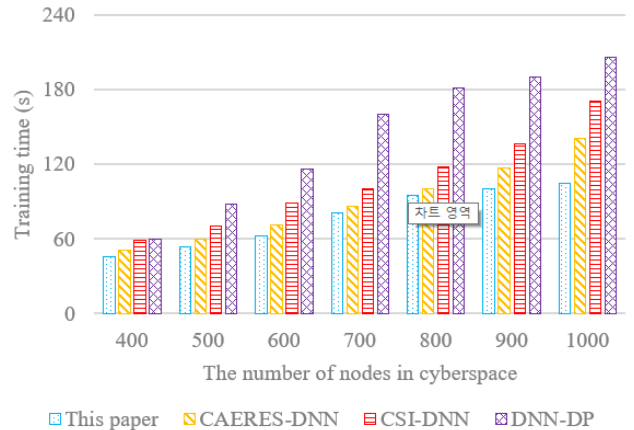


Fig. 7. The comparison of training time.

LLE is used to reduce the data dimensionality in order to avoid increasing the training time of high-dimensional data. In this paper, a certain number of samples are used to form a Mini-batch. It is obviously that the proposed method using Mini-batch has a good performance in running time, which is decreased by orders of magnitude. The Mini-batch used in this paper is a small part of training dataset. The data is divided into several groups, and parameters are updated according to the batch. In this way, data in batch jointly determines the direction of the gradient, so it is hard to deviate during descent and randomness is reduced. On the other hand, the number of batch samples is much smaller than the whole dataset, and the computation is not very large, which also reduces the running time (Fig. 8).

4.2.5. Recognition Accuracy

Data in complex-heterogeneous cyberspace mainly includes coordinates and other data such as application, IP and subject of network, if the data movement or data fault happens in data modeling, the selected model must recognize movement and fault in time. As shown in Fig. 9, VARMA model has a relatively average fault recognition accuracy, which is also the highest, because it is analyzed and modeled from both spatial and temporal attributes in complex-heterogeneous cyberspace. The recognition accuracy of ARMA model is relatively the lowest, but it performs well when the number of nodes ranges from 600 to 700. When the number of nodes is less than 800, the recognition accuracy of ARIMA model is better than that of ARMA model, but when the number of nodes is more than 800, the recognition accuracy of ARIMA Model is suddenly lower than 90%. While the recognition accuracy of ARCH model and GARCH model is stable at more than 90%.

4.2.6. Visualization Results

Additionally, we compare the visualization results using the GARR from the internet topology zoo. GARR is Italy's

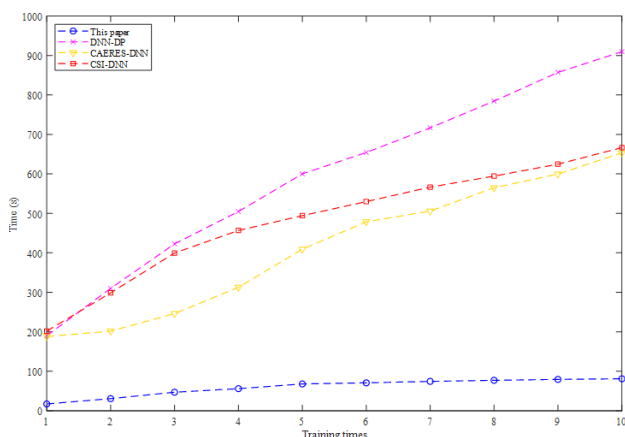


Fig. 8. The time comparison under different Mini-batches.

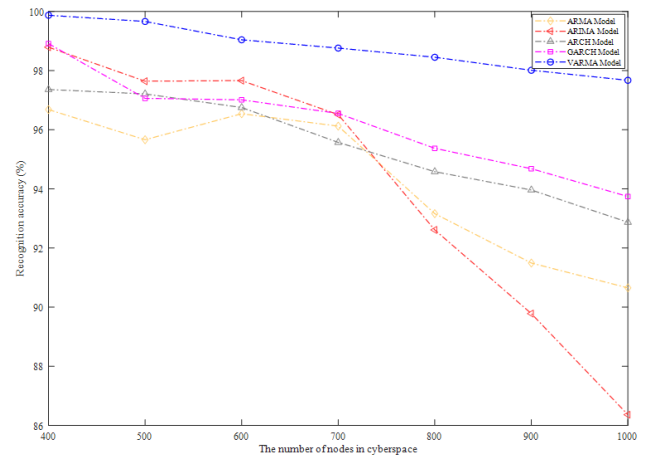


Fig. 9. The comparison of recognition accuracy with different models.

national research and education network, which stands for "Gruppo Armonizzazione Reti della Ricerca" (Italian for Research and Education Network Harmonization Group). It provides high-performance network connectivity and advanced services to the academic and research community in the country. GARR provides its users with a wide range of services, including high-speed internet connectivity, videoconferencing, cloud computing, virtual private networks, and access to e-learning platforms. Fig. 10 shows the visualization results of the four methods on GARR. The visualization result of DNN-DP, which exhibits no operations on nodes and edges, suggests a streamlined and concise representation. This can be advantageous for situations where

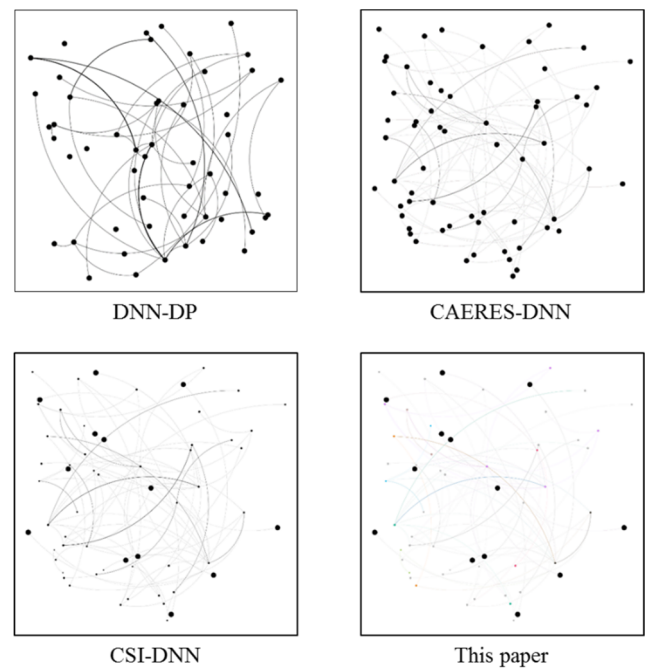


Fig. 10. The comparison of visualization results with different models.

simplicity and clarity are prioritized, allowing for a more focused understanding of the network's structure. In contrast, CAERES-DNN adjusts the weight of edges in its visualization result. This adjustment likely highlights the significance or relevance of certain connections within the network. By assigning different weights to the edges, the visualization can emphasize essential relationships and provide insights into the network's functional dynamics. Meanwhile, CSI-DNN ranks nodes by latitude in its visualization result. This arrangement based on latitude can introduce a geographical context, potentially aiding in interpreting network components and their spatial relationships. This approach might be beneficial in scenarios where geographic factors play a role, such as studying regional network connectivity or assessing network vulnerabilities across different locations. The method proposed in the paper achieves a relatively full visualization result, which effectively captures the complex heterogeneous nature of the GARR network. The proposed approach demonstrates a comprehensive representation that incorporates multiple aspects of the network's structure and behavior by utilizing techniques such as LLE for dimensionality reduction and vector ARMA for spatiotemporal data modeling. The visualization results presented in Fig. 10 highlight the diverse perspectives and visualization effects achieved by the different methods. By showcasing the unique contributions of the proposed method in achieving complex heterogeneous cyber cartographic visualization, the paper establishes its originality and potential value in the cyber security and safety applications field.

4.2.7. Dimensionality Reduction

This paper uses LLE as a dimensionality reduction technique for training cartographic coordinates data. In the following, we compare LLE with PCA, t-SNE, and Isomap regarding reconstruction error, neighborhood preservation, visualization quality, and computational efficiency. PCA is a linear technique that focuses on capturing the maximum

variance in the data, while t-SNE and Isomap aim to preserve the global and local structures, respectively. Reconstruction error measures how well the reduced-dimensional data can be reconstructed back to the original high-dimensional space. It quantifies the loss of information during the dimensionality reduction process. Lower reconstruction error indicates better preservation of the original data. Neighborhood preservation: Since LLE aims to preserve the data's local structure, evaluating the nearest neighbors' preservation is essential. Visual assessment becomes crucial if the purpose of dimensionality reduction is to facilitate visualization. Moreover, the computational complexity of the dimensionality reduction techniques should also be considered, especially for large-scale datasets.

Table 1 shows the comparison of dimensionality reduction. The reconstruction error for PCA was calculated as the sum of squared differences between the original and reconstructed data using the first two principal components. The neighborhood preservation score for t-SNE was calculated using the k-nearest neighbor graph with $k=10$. The reconstruction error for PCA was calculated as the sum of squared differences between the original and reconstructed data using the first two principal components. The neighborhood preservation score for t-SNE was calculated using the k-nearest neighbor graph with $k=10$. The reconstruction error for Isomap was calculated as the difference between the geodesic distances in the high-dimensional and low-dimensional spaces. The reconstruction error for LLE was calculated as the sum of squared differences between the original and reconstructed data in the high-dimensional space. In this example, LLE outperformed PCA regarding reconstruction error and neighborhood preservation, consistent with its ability to capture nonlinear relationships and preserve the local structure. Isomap also performed well regarding reconstruction error and neighborhood preservation, but LLE needed to be more computationally efficient. t-SNE strongly preserved the local neighborhood structure but did not have an explicit reconstruction error. In terms of

Table 1. Comparison of dimensionality reduction.

Technique	Reconstruction error	Neighborhood preservation	Visualization quality	Computational efficiency
PCA	0.21	Does not explicitly preserve neighborhood structure	May not capture nonlinear structure	Linear, computationally efficient
t-SNE	N/A	Strongly preserves local neighborhood structure	Can produce high-quality visualizations	Nonlinear, computationally expensive
Isomap	0.1238	Strongly preserves global structure and can capture complex nonlinear structure	Can produce high-quality visualizations	Nonlinear, computationally expensive
LLE	0.0523	Strongly preserves local neighborhood structure and can capture complex nonlinear structure	Can produce high-quality visualizations	Nonlinear, computationally efficient

visualization quality, all techniques can produce high-quality visualizations depending on the specific goals and characteristics of the data. Therefore, visual inspection of the plots may also be an essential metric for evaluating the effectiveness of the different techniques. By evaluating these metrics, it can be determined that LLE is suitable for complex heterogeneous cyber cartographic visualization.

V. CONCLUSION

This paper studies complex-heterogeneous cyberspace cartographic visualization. At first, we use LLE to reduce the data dimensionality. Then, a certain number of data samples are used to form a Mini-batch, and data after dimensionality reduction is trained in DNN. At last, in terms of temporal and spatial, we design the data model in order to realize cartographic visualization. Furthermore, the proposed method is simulated based on cyberspace datasets available on data.world, and the comparison experiments demonstrate that the proposed method is outperforming in precision ratio, recall ratio, F1 score, training time and recognition accuracy.

Although the proposed method has a good performance in training time comparing with the baselines. However, it is generally slow to train large DNN networks. In the future, we can find a network that can accomplish similar tasks, and then use part of its shallow network and parameters to perform simple extraction of input features, that is, transfer learning, which can not only speed up the training speed, but also require less training data.

REFERENCES

- [1] I. Priyadarshini and C. Cotton, "Intelligence in cyberspace: The road to cyber singularity," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 33, no. 4, pp. 683-717, 2021.
- [2] X. Hong, G. Zhang, and D. Lu, "Control strategies for crowd emotional contagion coupling the virtual and physical cyberspace in emergencies," *IEEE Access*, vol. 8, pp. 37712-37726, 2020.
- [3] Y. Koike, T. Matsubayashi, and S. Kondo, "Approaches to cyber-physical interactions linking the real world and cyberspace," *NTT Technical Review*, vol. 18, no. 9, pp. 31-35, 2020.
- [4] V. M. Krundyshev, "Identification of cyber threats in networks of industrial internet of things based on neural network methods using memory," *Automatic Control and Computer Sciences*, vol. 54, no. 8, pp. 900-906, 2020.
- [5] Z. Fan, C. Tan, and X. Li, "A hierarchical method for assessing cyber security situation based on ontology and fuzzy cognitive maps," *International Journal of Information and Computer Security*, vol. 14, no. 3/4, pp. 242-262, 2021.
- [6] K. B. Gemlau, L. Kohler, and R. Ernst, "A platform programming paradigm for heterogeneous systems integration," *Proceedings of the IEEE*, vol. 109, no. 4, pp. 582-603, 2021.
- [7] R. Xu, X. Da, H. Hu, L. Ni, and Y. Pan, "A secure hybrid satellite-terrestrial communication network with AF/DF and relay selection," *IEEE Access*, vol. 7, pp. 171980-171994, 2019.
- [8] F. Muller-Karger, M. Kavanaugh, K. Iken, E. Montes, F. Chavez, and H. Ruhl, et al., "Marine Life 2030: Forecasting changes to ocean biodiversity to inform decision-making: A critical role for the marine biodiversity observation network (MBON)," *Marine Technology Society Journal*, vol. 55, no. 3, pp. 84-85, 2021.
- [9] C. Lambert, M. Nahon, D. Charlmers, and G. Gilardi, "Cable control of an aerostat platform: Experimental results and model validation," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 2, pp. 620-628, 2007.
- [10] S. Jung and J. Choi, "End-to-end reliability of satellite communication network systems," *IEEE Systems Journal*, vol. 15, no. 1, pp. 791-801, 2021.
- [11] J. Zhang, J. Pang, and Z. Zhang, "Heterogeneity quantization method of cyberspace security system based on dissimilar redundancy structure," *Dianzi Yu Xinxu Xuebao/Journal of Electronics and Information Technology*, vol. 41, pp. 1594-1600, 2019.
- [12] M. A. Elliott, C. Nothelfer, C. Xiong, and D. A. Szafir, "A design space of vision science methods for visualization research," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1117-1127, Feb. 2021.
- [13] M. Li, F. Choudhury, Z. Bao, H. Samet, and T. Sellis, "ConcaveCubes: Supporting cluster-based geographical visualization in large data scale," *Computer Graphics Forum*, vol. 37, no. 3, pp. 217-228, 2018.
- [14] P. Kunz and J. Bobrich, "Multiscale cartographic visualization of harmonized datasets," *International Journal of Cartography*, vol. 5, no. 2-3, pp. 178-194, 2019.
- [15] C. Franke and J. Schweikart, "Mental representation of landmarks on maps: Investigating cartographic visualization methods with eye tracking technology," *Spatial Cognition and Computation*, vol. 17, no. 1-2, pp. 20-38, 2017.
- [16] J. Bac-Bronowicz and P. Grzempowski, "Regionalization of geographical space according to selected topographic factors in reference to spatial distribution of precipitation: Application of artificial neural networks

- in GIS," *Environmental Earth Sciences*, vol. 77, no. 18, p. 631, 2018.
- [17] M. Tsou, "Research challenges and opportunities in mapping social media and big data," *Cartography and Geographic Information Science*, vol. 42, no. sup1, pp. 70-74, 2015.
- [18] G. Ortiz-Jimenez, A. Modas, S. M. Moosavi-Dezfooli, and P. Frossard, "Optimism in the face of adversity: Understanding and improving deep learning through adversarial robustness," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 635-659, 2021.
- [19] J. Xu, H. Guo, H. W. Shen, M. Raj, X. Wang, and X. Xu, et al., "Asynchronous and load-balanced union-find for distributed and parallel scientific data visualization and analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 6, pp. 2808-2820, Jun. 2021.
- [20] M. Zhang, L. Chen, Q. Li, X. Yuan, and J. Yong, "Uncertainty-oriented ensemble data visualization and exploration using variable spatial spreading," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1808-1818, Feb. 2021.
- [21] J. Piazzentin Ono, J. Freire, and C. T. Silva, "Interactive data visualization in jupyter notebooks," *Computing in Science & Engineering*, vol. 23, no. 2, pp. 99-106, 2021.
- [22] J. D. Walker, B. H. Letcher, K. D. Rodgers, C. C. Muhlfeld, and V. S. D'Angelo, "An interactive data visualization framework for exploring geospatial environmental datasets and model predictions," *Water (Switzerland)*, vol. 12, no. 10, pp. 1-20, 2020.
- [23] S. Yoo, S. Jeong, and Y. Jang, "Gaze behavior effect on gaze data visualization at different abstraction levels," *Sensors*, vol. 21, no. 14, p. 4686, Jul. 2021.
- [24] A. Narayan, B. Berger, and H. Cho, "Assessing single-cell transcriptomic variability through density-preserving data visualization," *Nature Biotechnology*, vol. 39, no. 6, pp. 765-774, Jun. 2021.
- [25] A. Galletta, L. Carnevale, A. Bramanti, and M. Fazio, "An innovative methodology for big data visualization for telemedicine," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 490-497, 2019.
- [26] Y. Cheng, X. Wang, and Y. Xia, "Supervised t-distributed stochastic neighbor embedding for data visualization and classification," *INFORMS Journal on Computing*, vol. 33, no. 2, pp. 419-835, 2021.
- [27] G. Huang and H. Qu, "Data visualization and data fusion on the visual performance of illustration," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 6, pp. 8795-8803, 2020.
- [28] J. Wang, S. Zhuang, C. Miao, and C. An, "Model and application of cyberspace information system," *Tongxin Xuebao/Journal on Communications*, vol. 41, pp. 74-83, 2021.
- [29] J. Youn, H. Oh, J. Kang, and Shin D, "Research on cyber IPB visualization method based on BGP archive data for cyber situation awareness," *KSH Transactions on Internet and Information Systems*, vol. 15, no. 2, pp. 749-766, 2020.
- [30] Y. Zhang, G. Si, and Y. Wang, "Design and implementation of cyberspace war situation visualization system for joint operations," *Journal of Zhengzhou University (Engineering Science)*, vol. 39, pp. 45-51, 2018.
- [31] Y. Wang, S. Li, and L. Ren, "Automatic generalization methods of cyberspace point cluster features considering characteristics," *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University*, vol. 46, pp. 427-433, 2021.
- [32] J. Du, F. Ao, P. Li, and H. Ma, "Characteristics and knowledge representation of cyberspace situation information," *Journal of Data Acquisition and Processing*, vol. 34, pp. 500-508, 2019.
- [33] G. U. O. Qiquan, G. A. O. Chundong, H. A. O. Mengmeng, and J. I. A. N. G. Dong, "Develop visualization technology of cyberspace to support construction of comprehensive prevention and control system of cyber security," *Bulletin of the Chinese Academy of Sciences*, vol. 35, pp. 917-924, 2019.
- [34] Y. Wang, S. Li, X. Zhang, C. Zhang, and R. Wang, "Visualization of cyberspace information based on composite distance cartogram," *Journal of Information Engineering University*, vol. 21, no. 3, pp. 334-339, 360.
- [35] B. Jiang, G. Wan, and R. Xu, "Research on cyberspace division and visualization method," *Journal of System Simulation*, vol. 29, pp. 1-8, 2017.
- [36] X. Li, F. Yang, L. N. Wang, X. K. Yu, T. Fei, and N. Jiang, "A survey of mapping methods for cyberspace," *Journal of Geomatics Science and Technology*, vol. 36, pp. 620-626, 631, 2019.
- [37] Y. Wang, M. Gu, J. Ma, and Q. Jin, "DNN-DP: Differential privacy enabled deep neural network learning framework for sensitive crowdsourcing data," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 1, pp. 215-224, 2020.
- [38] P. Ferrand, A. Decurninge, and M. Guillaud, "DNN-based localization from channel estimates: Feature design and experimental results," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*, Taipei, Taiwan, 2020, pp. 1-6.
- [39] M. Li, Y. Jiang, Z. Wu, and R. Fan, "Real-time prediction of smoke spread affected by multiple factors in subway tunnel using CAERES-DNN model," *Fire Technology*, vol. 57, no. 4, pp. 2025-2059, 2021.
- [40] R. Qin, X. Fu, J. Dong, and W. Jiang, "A semi-greedy

neural network CAE-HL-CNN for SAR target recognition with limited training data," *International Journal of Remote Sensing*, vol. 41, no. 20, pp. 7889-7911, 2020.

- [41] C. Yang, W. Jiang, and Z. Guo, "Time series data classification based on dual path CNN-RNN cascade network," *IEEE Access*, vol. 7, pp. 155304-155312, 2019.

AUTHORS



Lan Zhang is currently pursuing her Ph.D. at Information Engineering University, Zhengzhou, China. Also, she is an engineer at the major of surveying and mapping. Her research interests include digital map, deep learning, etc.



Zhiyong Liu is currently pursuing his Ph.D. at Wuhan University, Wuhan, China. Also, he is an engineer at the major of resource and environment. His research interests include satellite remote sensing, deep learning, etc.



Guangxia Wang is a full professor at Information Engineering University, Zhengzhou, China. Her research interests include digital map, AI, etc.



Lei Xiong is currently pursuing his master degree at Wuhan University, Wuhan, China. Also, he is an assistant engineer at the major of resource and environment. His research interests include satellite remote sensing, geographic information system, deep learning, etc.

