

## **Research Article**

# In-Depth Analysis of Railway and Company Evolution of Yangtze River Delta with Deep Learning

## Renzhou Gui 🕞, Tongjie Chen 🕞, and Han Nie 🕞

The Department of Electronic and Communication Engineering, Tongji University, Shanghai 201804, China

Correspondence should be addressed to Renzhou Gui; guishq16@lzu.edu.cn

Received 24 August 2019; Revised 27 November 2019; Accepted 23 December 2019; Published 20 January 2020

Academic Editor: Ludovico Minati

Copyright © 2020 Renzhou Gui et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The coordinated development of smart cities has become the goal of world urban development, and the railway network plays an important role in this progress. This paper proposes a solution that integrates data acquisition, storage, GIS visualization, deep learning, and statistical correlation analysis to deeply analyze the distribution data of companies collected in the past 40 years in the Yangtze River Delta. Through deep learning, we predict the spatial distribution of the company after the opening of the train stations. Through statistical and correlation analysis of the company's registered capital and quantity, the urban development relationship under the influence of the opening of the railway is explored. Going forward, the use and application of such analysis can be tested for use and application in the context of other smart cities for specific aspects or scale.

## 1. Introduction

Nowadays, the development of smart cities has become an emerging model and key target of world urban development [1]. IBM defines "smart cities" as the use of information and communication technologies to sense, analyze, and integrate key information about the city's core systems in operation [2]. It is predicted that more than 60% of the population will live in an urban environment [3] by 2030 and this record will further reach 68% by 2050 [4]. With the notion and the theory of "urban planning" and "smart city" becoming more complex [5], much more advanced technical support are required [6]. At present, the data-centric [7, 8] "smart city" is still in its infancy, and its development is challenged by noneffective process of huge amount of data and restricted extraction of data values [9]. New technologies such as big data [6, 10, 11] processing, data mining, artificial intelligence, and deep learning are important tools to build smart cities [6, 12]. However, the analysis of geospatial data in many developed cities ignores some important social and economic management elements. That is, there is a disconnect and asymmetry among information collection, acquisition, and application. And the coverage and utilization rate at the level of time and space is not comprehensive enough to meet the needs of government management [13].

As one of the most important economic entity in the world, China has attracted worldwide attention. Over these years, China has been vigorously promoting its urbanization strategy and strived to achieve the world's highest urbanization growth rate and broad development space [14]. And as an important active economic center of China, the urban agglomeration in the Yangtze River Delta [15-17] has formed five levels of urban scale hierarchy, showing the characteristics of "pagoda type" [18] that are mutually inclusive, mutually integrated, and mutually infiltrated. The in-depth analysis of the companies and economic data in the Yangtze River Delta makes sense in many aspects: It is conducive to the upgrading of new technologies and the mastery of the laws of urban economic development. That is to say, with the continuous updating of technology, this analysis process can help to reveal the inherent laws of economic development [19, 20] of human beings.

The purpose of this study is to design and develop a solution that combines the long-term company registration information in the Yangtze River Delta and new technologies to enable in-depth analysis of the relationship between railways and the company evolution. To support this research, company registration information (such as spatial location and opening time) of nearly forty years in the Yangtze River Delta were collected from the Internet. Big data, deep convolutional neural networks, and data correlation analysis techniques were combined and used to achieve effective storage, visualization, in-depth analysis and prediction of the mutual promotion, and development of railway and company. And a convolution-deconvolution hybrid deep learning network is proposed in this paper to process the spatial and temporal data of 314 railway stations in the Yangtze River Delta. By making full use of these data and tools, we have carried out research on related questions: the contribution of different regions (e.g., Shanghai and Suzhou) to the development of the entire Yangtze River Delta, the interaction and influence among the development of different regions, the role of different railways in the development of the surrounding and urban areas, and make it possible to predict spatial distributions of the company after the opening of the train stations through the deep learning method. This technical solution can be served as an effective method to solve the complex problems associated with smart cities.

This work is organized as follows. Section 2 shows related works. Section 3 discusses the acquisition and storage methods of network big data, the data analysis network model of deep learning, the data visualization, and the theoretical process related algorithms. Section 4 presents the results of visualization, deep learning, and relevant analysis. Section 5 discusses the results in Section 4. Finally, Section 6 summarizes the contributions and presents some future research directions.

## 2. Related Work

There are many factors affecting the economic development of urban agglomerations. Railway construction is one of the important factors and has a great impact in the Yangtze River Delta [21]. With its enormous social and economic benefits, it has become an important mean to develop urban economy and improve the industrial structure and the living standards of residents [22]. The existence of the transportation system is one of the essential differences between survival system and market economy [23, 24]. There have been many studies on the data analysis and application of railway transportation, economy, and smart cities. Some of the literature only studies intelligent transportation [25] and intelligent economic development [26] or explores the impact of transportation and economic development [27]. However, there are still some deficiencies in the overall indepth analysis of these three aspects by using the emerging deep learning and big data tools.

Smart city research is data centric, and the entire process includes data acquisition, collation, storage, processing, visualization, and analysis. Getting the data for smart city is the first of these. In recent years, big data on the Internet has become an important source for data analysis, such as the various social medias, Bulletin Board System (BBS), news, and other websites [28]. Big data is a term for massive data sets having large, more varied, and complex structures with the difficulties of storing, analyzing and visualizing for further processes or results [10]. These data have the characteristics of large quantity, wide range, and high value and have made great contributions to the financial, economic, resource, energy, and tourism industries required for urban development [25, 27, 29, 30]. Sen et al [31] used the Google location data from around the world to collect walking information from 25 cities on five continents. It also showed the useful applications of these data in urban analysis, such as how different areas of the city compare in terms of accessibility and which areas of the city benefit the least from the least intervention. Alberto et al [32] used web tools to automate data collection, collation, and visualization of complex healthcare research, including social network analysis.

Once the data have been acquired and organized, relevant tools can be used for initial visualization and analysis. Afaneh and Shahrour [33] introduced the use of geographic information for data management and visualization in the Smart City Project (SunRise, Smart City) at the Lille University Science Park. It used smart sensors to collect operational data of water and energy networks and used Geographic Information System (GIS) to manage and visualize assets and operational data of smart city projects. Yamamura et al [34] proposed a "GIS-BIM"-based urban energy planning system that includes GIS database construction and analysis to obtain optimized technical and policy solutions to realign urban infrastructure. However, the management and visualization of these data is too dependent on the software itself, such as ArcGIS. Overreliance on these software, data interface, and algorithm modification optimization may be problematic, resulting in an internal mechanism that cannot flexibly and efficiently explore data.

For in-depth learning and mining of big data, especially geographic data, today's machine learning (ML) [35] technology has played a key role in this work [36]. ML is a discipline focused on two interrelated questions. The ML algorithms can be viewed as searching through a large space of candidate programs, guided by training experience, to find a program that optimizes the performance metric [37]. It is able to extract as much useful information as possible and gain new insights from these data, simulations, and interaction between the two. Tarutani et al. [38] applied ML regression model techniques to predict temperature distribution in the data center, reducing air conditioning power consumption by approximately 30%. Chin et al. [39] studied four well-known ML classification algorithms (Bayesian Network (BN), Naive Bayes (NB), J48 Tree Classifier, and Nearest Neighbor Algorithm (NN)) in correlating weather data (in particular, the impact of rainfall and temperature on the short-distance travel of London cyclists). Through the evaluation in terms of accuracy, credibility, and speed, the conclusion is that the decision tree J48 algorithm performs best in accuracy, while the kNN IBK algorithm builds the model fastest.

Compared with the traditional ML method, deep learning can deeply study the diverse structural features in high-dimensional data and explore the class distribution probability of data [40]. Deep learning methods are the multilevel representation learning method with representations, which is obtained by combining simple but nonlinear modules, each of which converts a level (from the raw input) into a higher, slightly more than the abstract level [41, 42]. Balchandani et al. [43] proposed a deep learning framework to analyze street photographs and to detect if the streets were dirty by detecting garbage objects. The framework better analyzes the automation framework for street cleaning problems by analyzing, finding, and efficiently scheduling cleaners in areas that require more attention. Dongjie et al. [44] realized a detection of the high-accuracy locators based on the Faster R-CNN model, which has a high engineering application value. They also used the Hough transform to detect the skeleton contour of the locator and filtered the optimal fitting straight line of the locator through the filter line mechanism, which performed a noncontact accurate measurement of the slope of the positioner. Camero et al. [45] proposed a new technology based on the deep learning and regression neural network to solve the problem of parking occupancy prediction. The occupancy rate of 29 parking lots in Birmingham, England, was tested in the data of 11 weeks. The results show that the method is accurate to the way people use in their daily lives and is superior to existing competitors.

In summary, we find that the Internet has huge and valuable data and need to be acquired and stored by using effective tools. Faced with such huge data, traditional relational databases cannot satisfy large-scale and high-concurrency dynamic queries nor can they effectively support deep learning and subsequent data research. At present, big data research in smart cities is still in its infancy. Traditional artificial intelligence algorithms (such as neural network, genetic algorithm, artificial bee colony, particle swarm optimization, cuckoo search algorithm, flower pollination algorithm, chicken swarm optimization, and bat algorithm) can still theoretically be used, but are heavily limited to the calculation of small data sets, not for smart city big data analysis [46]. And the current research progress has not produced enough value and contribution [47]. The effective visualization and processing and effective deep learning and analysis of big data need to be organically combined.

#### 3. Materials and Methods

The implementation process of this paper is shown in Figure 1. First, the Internet resources were fully utilized to obtain the established time, location, latitude and longitude, type and registered capital of all companies registered in the Yangtze River Delta from 1980 to 2018, and the related opening time of 314 railway stations (15 railways). Next, by developing GIS software, the data are displayed on the map. After that, according to the opening time and location of each railway line, the company registration data are screened to calculate the distribution density matrix of the company. These density matrices are imported into the design's deep learning model. Through this model, the distribution forecasting training for the three years after the opening of the railway will be carried out to find out the in-depth impact of the railway on the economy. The historical data of the

3.1. Data Acquisition. This paper uses Scrapy [48] to develop a distributed framework for website data acquisition. These data are from Enterprise Yellow Pages of China which include base information of company we need and are free and public to society. Thus, there are no legal and ethical issues [49]. By analyzing the characteristics and locations in the website, the data are efficiently collected in a breadth-first manner. The latitude and longitude conversion of the location information is realized by combining the Baidu map API.

The regular expression [50] method is used when actually extracting page data. However, when extracting the data required by the webpage, only using this method cannot flexibly identify the data features based on the page content. Moreover, due to the different content of different web pages (especially when there are similar tags in the code of web pages), it is easy to grab the wrong data of the wrong location, or even unable to grab any information correctly. According to the tree structure characteristics of the webpage, this paper analyzes the subpages and the nodes of data in the tree structure of the webpage and uses the XPath [51] method to extract the relative path of the data. This method can extract the corresponding data accurately and efficiently, and then finely process the obtained data by the method of regular expression. Finally, the data are stored and saved as a JSON file for importing into the database.

3.2. Data Storage. Relational databases are slow to respond to large and highly concurrent dynamic queries. And its storage method is also inefficient when facing big data. Through analysis, it is found that there is no other relationship between the data in these unstructured data in the Yangtze River Delta. But the amount of data is large, and there is a need to better support a large amount of quick access to tools such as GIS. For the use of these data, the platform needs to store once and query multiple times without further processing and update operations on the data. In the Internet field, NoSQL database [52, 53] technology based on nonrelational database has been successfully applied. It breaks the relational model of traditional databases, making data stored in a more free way, without relying on the relationship between fixed table structures and data. This is great for storing big data.

NoSQL databases are divided into key-value databases, columnar databases, document databases, and graph databases [54]. By comparing the advantages and disadvantages of each database, this paper finally adopts the document-oriented MongoDB database [55, 56], a database based on distributed file storage. The query language of MongoDB is very powerful, and it uses binary json (bjson) form of the document for storage and can store more complex data types. This database is also easy to be deployed and used. Li et al. [56] proposed a hybrid distributed storage strategy for spatial big data, in which MongoDB database was used as the



FIGURE 1: Flow chart for collection, processing, deep learning, and analysis of enterprise data in the Yangtze River Delta.

engine of distributed spatial data to meet the needs of big data storage and the needs of traditional GIS applications.

This paper stores two types of data, which are company data and train station data. For the company data, we acquired 1,921,973 data, including all kinds of companies in the Yangtze River Delta region of China for 40 years. Each data includes the company's completion time, the registered capital, the type, the type of product manufactured, the detailed address, and the latitude and longitude corresponding to the company's address. And for the train station data, the information of the Yangtze River Delta railways including 314 railway stations (15 railways). These railways are including Ningqi railway, Xinchang railway, Hening railway, Suhuai railway, Shanghai-Kunming railway, Shanghai-Nanjing railway, Nanjing-Hangzhou high-speed rail, Xiaoyu railway, Hangzhou-Ningbo high-speed rail, Beijing-Shanghai high-speed rail, and Beijing-Shanghai railway.

3.3. Data Visualization. Data visualization plays a very important role in smart cities [57]. A common visualization tool is GIS, which can display data in a variety of visual ways in the corresponding map. However, relying solely on dedicated GIS software, it is difficult to achieve more free data calculation and analysis. Therefore, when writing related processing programs, we use ArcGIS Engine embedded in C# programming to construct a special GIS application to help us realize spatial visualization of data, as shown in Figure 2. The interface implements a variety of free control modes and can adjust the display time range, regional range, and company type range by the user and also perform regional statistics.

We optimized the calculation and clustering of our data based on the k-means [58, 59] clustering algorithm and the simple linear iterative clustering (SLIC) [60] algorithm so that the clustering algorithm can balance the spatial position of the data and the registered capital of each company. Interestingly, the results of the development of clusters are indeed the current shape of "three provinces and one city" (Jiangsu Province, Anhui Province, Zhejiang Province, and Shanghai), while "one city" area occupies more regions than the real Shanghai. As the economic center of the Yangtze River Delta, Shanghai has geographically driven the entire Yangtze River Delta to develop its economic and business [61].

3.4. Data Preparation. By comparing the time of the completion of the various railway lines and the geographical range of  $40,000 \times 40,000$  m<sup>2</sup> centered on the railway station, we used the built visualization application to extract the railway line before and after the completion of the railway line. By dividing the spatial data into  $200 \times 200$  space, we can get the company's 2D data and density matrix for each site.

In order to prevent overfitting (overfitting [62] is a modeling error which occurs when a function is too closely fit to a limited set of data points and can be alleviated by increasing the effective sample size.) of the data and increase the sample size, we performed symmetry and rotation on the obtained density data matrix, which increased the sample size by 6 times, and finally obtained 2198 (314 \* 7) data. According to 4:1, the data is divided into 1758 (2198 \* 0.8) training data and 440 (2198 \* 0.2) test data sets. Since the input is the time-space distribution status information of the companies accumulated for 10 years or 20 years before the

## Complexity



(a)



FIGURE 2: Data visualization interface designed in this paper. (a) Spatial distribution map of all companies collected from 2008 to 2018. (b) A result graph of clustering company data of spatial distribution using the improved *k*-means algorithm (the number of clusters is chosen to be 5).

completion of the railway station, the information automatically carries the road traffic information, the prosperity level information, and the government policy support information around the railway.

3.5. Deep Learning Network. Deep learning network is an extension of the common artificial neural network, with many hidden layers, so it has powerful nonlinear mapping ability [41], as shown in Figure 3. It can perform independent deep learning and feature extraction on data, thus achieving good model performance [63]. Among them, deconvolution is a process of convolution inversion, which is widely used in image restoration [64] and generation to recover clear images from related data [65]. Moreover, since the size of the intermediate layer image is made smaller by using the convolutional layer, the deconvolution restores the original size of the image and even makes it larger.

In this paper, a new deep neural network model consisting of convolution and deconvolution layers is designed, as shown in Figure 4. It is used to learn and predict the density matrix of the  $40,000 \times 40,000$  m<sup>^</sup>2 area centered on the railway station before and after the opening of railway lines. The input data is the existing company quantity density matrix and registered capital density matrix before the railway station is built. The output tag data are the newly added company quantity density matrix after the site is opened for three years.

The model consists of 3 convolutional layers and 3 deconvolutional layers. The learning rate is 0.0001 and the batch size is 16. The activation function of each layer is a Relu function. Drawing on the idea of Residual Networks [66], the output of the second layer and the output of the fourth layer are added together, calculated by the Relu function, and then sent to the fifth layer deconvolution layer. This makes it easier to pass the error of the loss function to the deep layers of the network, helping the network accelerate training. At the same time, the regularization function is added to the loss function to prevent overfitting, and the GPU is used to speed up the operation.

3.6. Loss Function. In the field of deep learning, the minimum mean square error (MMSE) [67] and the cross entropy [68] function are known commonly, used as loss functions. However, they are both used to train the network only with the numerical difference as the loss value, so it lacks the characteristics of the spatial distribution of data in two dimensions. As shown in Figure 5, assuming the true distribution as shown in Figure 5(a), the two possible prediction distributions obtained from the deep learning network are Figures 5(b) and 5(c). If only the numerical difference is calculated, then the loss value obtained by subtracting Figure 5(a) from Figure 5(b) will be the same as the loss value obtained by subtracting Figure 5(a) from Figure 5(c). But the real situation is that Figure 5(b) is closer to the actual distribution of Figure 5(a), so the loss value from Figure 5(b) should be smaller than the one from Figure 5(c). Therefore, considering the number and location



FIGURE 3: Deep neural network model.

characteristics of the two-dimensional spatial distribution of data, this paper attempts the loss function form formed by various combinations of the MMSE, Two-Dimensional Discrete Fourier Transform (2D-DFT), and Cosine Similarity [69].

MMSE can better constrain the number of predicted density map matrices. 2D-DFT can help to better constrain the distribution of predicted data and even achieve image reconstruction [70]. Cosine similarity can better constrain the similarity of distribution [71] and measure the similarity of different objects to achieve image classification [72]. Considering that the loss of the MMSE is too small due to the normalization of the input data, we do not divide the sum of the squares of the differences by the number of elements of the entire matrix, but only divided by the slide length of the matrix. For the cosine similarity formula  $\cos\langle x, y \rangle =$  $(\sum x \cdot y/x \times y)$ , the value distribution is [-1, 1], where 1 represents the largest correlation and -1 means the lowest correlation. We map the value of the similarity to 0 to 2 (0 represents the maximum correlation and the lowest loss value) to achieve network training. The formulas for each loss function and the total loss function formula are as follows:

$$\text{Loss}_{\text{MMSE}} = E(y-x)^2 = \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=0}^{n-1} \left( y_{ij} - x_{ij} \right)^2, \tag{1}$$

$$fft2(x)_{lk} = \frac{1}{M \times N} \sum_{u=0}^{M-1} \sum_{\nu=0}^{N-1} x_{u\nu} \times e^{j2\pi((ul/M) + (\nu k/N))}, \qquad (2)$$

$$Loss_{2DDFT} = E(fft2(y) - fft2(x))^{2}$$
  
=  $\frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=0}^{n-1} (abs(fft2(y)_{ij}) - abs(fft2(x)_{ij}))^{2},$   
(3)

$$Loss_{Cosin} = 1 - \frac{\sum x \cdot y}{x \times y}$$
  
=  $1 - \frac{\sum_{i=1}^{n-1} \sum_{j=0}^{n-1} y_{ij} \times x_{ij}}{\left(\sum_{i=1}^{n} \sum_{j=1}^{n} \sqrt{x_{ij}^{2}}\right) \times \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \sqrt{y_{ij}^{2}}\right)},$  (4)



FIGURE 4: Proposed deep neural network model.



FIGURE 5: Several possible distributions of two-dimensional data. (a) True distribution. (b) One possible prediction distribution 1. (c) One possible prediction distribution 2.

$$Loss_{MC} = 0.6 \times Loss_{MMSE} + 0.4 \times Loss_{Cosin},$$
 (5)

$$Loss_{MF} = 0.6 \times Loss_{MMSE} + 0.4 \times Loss_{2DDFT}, \qquad (6)$$

$$\text{Loss}_{\text{CF}} = 0.6 \times \text{Loss}_{\text{Cosin}} + 0.4 \times \text{Loss}_{\text{2DDFT}},$$
 (7)

$$Loss_{MCF} = 0.4 \times Loss_{MMSE} + 0.4 \times Loss_{Cosin} + 0.2 \times Loss_{2DDFT},$$
(8)

where equation (1) is the loss function formula of MMSE and n is the side length of the density matrix. Equation (2) is the Fourier transform formula of the density two-dimensional matrix, and equation (3) is the loss function of 2D-DFT. Equation (4) is the loss function formula of cosine similarity, equation (5) MC is the loss function formula of the combination of MMSE and cosine similarity, and equation (6) MF is the loss function formula of the combination of the combination of the combination of cosine similarity and 2D-DFT. Equation (7) CF is a loss function formula of the combination of the combination of the combination of the three formulas. There are 7 loss functions in total, and all weight coefficients are obtained through multiple trials.

3.7. Data Correlation Analysis. Correlation coefficients can be used to discover correlations between data and help

analyze data such as economics [73], transportation [74, 75], and life [76] to serve humans. By obtaining new data for each year before and after the completion of the railway, we can use the Pearson coefficient to obtain the degree of development and correlation of each region. The Pearson coefficient is calculated as follows:

$$r_{ij} = \frac{\sum (x_i - \overline{x_i}) (y_j - \overline{y_j})}{\sqrt{\sum (x_i - \overline{x_i})^2} \sqrt{\sum (y_j - \overline{y_j})^2}},$$
(9)

where  $x_i$  and  $y_i$  represent the sequence of increasing number of years in the *i*th and *j*th region,  $\overline{x_i}$  and  $\overline{y_i}$  represent the mean of the sequence, and  $r_{ij}$  is the correlation of the two regions. The correlation coefficient ranges from -1 to 1, where +1 indicates positive correlation and -1 indicates negative correlation.

However, simply calculating the value of this increase cannot fully reflect the change in the number of each region. Since the number of companies growing in each region is always positive or zero, the discrimination may not be obvious enough when comparing the correlation coefficient values between different regions. Therefore, this paper further combines the change of the data growth value of each region (that is, the first derivative of the growth value) and weights the new correlation coefficient matrix. The new correlation coefficient calculation formula is as follows:



FIGURE 6: Deep learning training loss function curve corresponding to different combinations of loss functions. (a) MMSE. (b) Cosine Similarity. (c) 2D-DFT. (d) MMSE-2D-DFT. (e) Cosine Similarity-2D-DFT. (f) MMSE-cosine Similarity. (g) MMSE-cosine Similarity-2D-DFT.

$$r_{ij}' = \frac{\left(\left(\sum(x_i - \overline{x_i})(y_j - \overline{y_j})/\sqrt{\sum(x_i - \overline{x_i})^2}\sqrt{\sum(y_j - \overline{y_j})^2}\right) + \left(\sum(x_i' - \overline{x_i'})(y_j' - \overline{y_j'})/\sqrt{\sum(x_i' - \overline{x_i'})^2}\sqrt{\sum(y_j' - \overline{y_j'})^2}\right)\right)}{2}, \quad (10)$$

where  $x'_i$  represents the first derivative of the original data x,  $x'_i = x_{i+1} - x_i$ .

## 4. Results

4.1. Deep Learning. The seven different loss functions are compared by using them in the deep learning network model designed in this paper. As shown in Figure 6, we finished the training separately and evaluated the test set. The loss function value curve of these seven loss function trainings is shown in Figure 6. The red curve is the loss value of the training data, the green curve is the training curve of the test set, and the abscissa unit is the size of each batch (value: 16).

It can be seen that these loss functions with the MMSE are trained well to reach the best value. The loss value of the test set also decreases as the number of training increases, and the final trend is flat. The training of the deep neural network seems to be very successful. But the loss function of 2D-DFT has the worst training effect and cannot even train the net. The loss function of cosine similarity can help the training, but the effect is poor. And the loss value changes are more jittery and not stable enough. Therefore, MMSE is the best loss function that helps the network to effectively train among them.

However, as mentioned before, MMSE can only constrain the similarity of the result values and cannot achieve

effective prediction of the distribution. Table 1 shows three examples of test sets. We will have the MMSE value (which is different from the loss function, and the result is the average of the number of matrix elements) and the cosine similarity value (unlike the loss function, the result is the similarity coefficient of [-1, 1], where 1 represents completely similar and -1 represents completely opposite). The first row is the quantity distribution of the railway before the completion, the second row is the distribution of the number of new companies after three years of construction, and the next 7 rows are, respectively, predicted by 7 different loss functions. For the convenience of the display, only the density matrix is displayed. In these figures, the yellower the color is, the lower the density value (the minimum is 0), and the redder the color, the greater the density value (the highest value we choose is 10. If the value is greater than 10, the color will be pure red). The forecast results can also reflect the possibility of presence in the region.

This paper also does a statistical average of the MMSE and cosine similarity of all test data sets, as shown in Figure 7.

4.2. Statistics and Correlation Analysis. In order to further analyze the number of companies before and after the completion of the railway station, the detailed statistics of

TABLE 1: Comparison of images, minimum mean square error, and cosine similarity of three prediction results.



Continued.
÷
LABLE



TABLE 1: Continued.



Continued.
÷
TABLE

Xinggongtang station in Xuanhang railway (1994/07)		0.14423271414878933 0.662771618719908282		0.000053113823069145 0.556173118171807028
Huishan railway station in Shanghai- Nanjing railway (2010/07)		0.7541079039032181 0.33967178745360267		0.000165248184589094 0.338455365847816467
Hefei railway station in Yuliu railway (2013/12)		4.620754636255468 0.597224169228503626		0.000364660292423330 0.577345932340722066
ttion in the regional center, which it belongs, and the	Image	MMSE value cosine similarity	Image	MMSE value cosine similarity
The name of the train sta the name of the line to v opening time	Cosine-2D-DFT loss		MMSE- cosine loss	





FIGURE 7: MMSE value and cosine similarity evaluation results of 7 kinds of loss function corresponding to test set. Unlike the loss function of the cosine similarity, the similarity value here is the similarity coefficient distributed in [-1, 1], where 1 represents completely similar and -1 represents completely opposite.



FIGURE 8: Statistical chart of the number of newly added companies three years after the opening of railway stations in three years. The abscissa indicates the number before completion of the railway, and the ordinate indicates the average number of the newly added company after 3 years of construction.

the company are shown in Figures 8 and 9 (the statistical value of 0 means that there is no data for this area during statistics).

From these figures, we can see that under the influence of the construction of the railway station, it is not that the area will have more new established companies when the area has more densely companies before the station is completed. The distribution results are bimodal, and they may even have a negative impact in a certain area.

Which region contributes the most to the development of the Yangtze River Delta region? We counted the data of Fuyang, Hefei, Nanjing, Zhenjiang, Suzhou, Shanghai, Lu'an, Wuxi, Xuancheng, Hangzhou, Shaoxing, Ningbo, Taizhou, Wenzhou, Lishui, Yiwu, Nanchang, Fuzhou, Yichun, Quzhou, Anqing, Changzhou, Suzhou, Huai'an, Yancheng, Xuzhou, and other places. As shown in Figure 10, there are 22 cities and their annual development changes. Shanghai is the fastest growing city in the entire Yangtze River Delta region. And Hangzhou, Wenzhou, and Nanjing are ranked behind Shanghai.

At the same time, we compare the three different correlation coefficient matrix diagrams mentioned above. As shown in Figure 11, the first row or the first column is the statistics of the entire Yangtze River Delta. By comparing the correlation coefficient matrix and the derivative correlation coefficient matrix, we can see several cities that differ greatly from the entire Yangtze River Delta region, such as Shanghai and Lishui. For the range of correlation coefficients, 0.8–1.0 means that there is a very strong correlation between the two places, 0.6–0.8 means that there is a strong correlation between the two places, 0.4–0.6 means that the two places have a moderate correlation, 0.2–0.4 means that the two places have a weak correlation, and 0.0–0.2 means that the two places have very weak correlation or no correlation.

In 2007, China's train speed up for the sixth time [77, 78], which has played a great role in promoting

#### Complexity



FIGURE 9: Statistical chart of the newly added registered capital scale of the company after the opening of railway stations in three years. The abscissa indicates the average value of the registered capital of the newly added company after the opening of the railway.



FIGURE 10: Chart of the growth of companies in the Yangtze River Delta from 1990 to 2018.

economic development and coordination and interaction between cities (verified the turning time of the curve change in Figure 10). Taking 2007 as the time division point, we calculated the correlation coefficient statistics for the 5 years before and 5 years after 2007 (the combined correlation coefficient statistics using the third combined derivative). As shown in Figure 12, through the comparison of coefficient matrices, it can be clearly found that the construction of railways has a great impact on the development and change of cities.

Similarly, since the Beijing-Shanghai high-speed rail was opened in 2011 and the Ningbo-Hangzhou high-speed rail was completed in 2013. As shown in Figure 13, we compared the five-year correlation coefficient matrix before and after 2013.

#### 5. Discussion

Through the application of new technologies, a convolutiondeconvolution hybrid deep learning network is adopted in this paper to process the spatial and temporal data of companies in 314 railway stations in the Yangtze River Delta. The network realized the prediction of density matrix and acquired the distribution around the railway station, as

#### Complexity









(b) FIGURE 11: Continued. 17



FIGURE 11: Matrix diagram of relevance coefficient (The data are a time series of the number of cities in the Yangtze River Delta from 1980 to 2018). (a) Correlation coefficient matrix of time series of the individual growth value. (b) Correlation coefficient matrix of time series of first-order derivative of the individual increment value. (c) Correlation coefficient matrix that weights the sum of (a) and (b).

shown in Figure 4. Through seven different types of loss functions, we get different training results of in-depth learning models (Figure 6). From Table 1 and Figure 8, we can see the importance of the distribution-related loss function for network training. In these loss functions, we can see that although the difference in MMSE loss values is small, the allocation results are poor and even the results of the company's densely distributed areas are lost. We also found that 2D-DFT also performed badly in the actual training results. The result of the training can only reflect the approximate distribution shape. When combined with other loss functions, the combination with MMSE can improve the similarity results, but still relatively low, and the combination with cosine similarity is worse. For the loss function of cosine similarity, although the individual training results are not very satisfactory, they can only grasp the shape and characteristics of the distribution. In the process of combining with MMSE, the network training is not only small in error but also more similar in distribution. The results are most consistent with the density matrix.

In the process of statistics of the newly added registered capital scale of the company after the railway station was built for three years (Figure 9), we found that the number of statistics shows a bimodal shape, which means that the economic scale of a certain city has grown linearly in the initial stage of the number. Company growth achieves the highest when the number before the completion of the project reached 45k to 50k, attaining an average of 24,489.33. We can regard this stage as the growth period. However, after this stage, the impact of the railway on it has rapidly decreased. After the number reaches 85k, it shows a leaping growth and then remains stable. We can regard this stage as the maturity stage of development. Comparing the statistical chart of registered capital in Figure 9, we find that the results



FIGURE 12: Continued.



FIGURE 12: Matrix diagram of relevance coefficient of corporate development of cities in the (a) five years before 2007 and the (b) five years after 2007.

are similar, but they are more prominent in the 45k to 50k part. And the average of their registered capital reached 1,944,758,200 yuan. Followed by 51.598 million yuan from 35k to 40k. The registered capitals added in other corresponding areas are small.

Finally, from the perspective of economic correlation, the temporal and spatial evolution of 23 regions (including 22 cities and the whole Yangtze River Delta) in the past 40 years is counted and analyzed (Figure 10). The result shows that the overall growth rate is relatively slow before 2007, but after that, the whole Yangtze River Delta has a rapid growth momentum (at this time, China has also carried out the sixth train speed increase [77, 78]), in which Shanghai is the fastest growing city in the whole Yangtze River Delta. The number of Shanghai companies has shown a leap in growth. This result is very similar to the GDP results of the National Bureau of Statistics of China [79]. It also confirms the leading position of Shanghai and the huge difference between Shanghai and the entire Yangtze River Delta (Figure 11).

Through further observation of the Figure 11, we found that before 2007, the development of each city was not only different but also poorly correlated with the whole. After 2007, although the new changes are great, the difference between cities gradually decreases. By comparing the first column of the matrix, each city gradually keeps up with the growth momentum of the whole Yangtze River Delta. This shows that the construction or acceleration of railways has a great correlation with the development and changes of the cities. It is found that, in the five years after 2013 (Figure 13), the relationships among the cities in the Yangtze River Delta have been further strengthened. Especially the relationships



(a) FIGURE 13: Continued.



FIGURE 13: Matrix diagram of relevance coefficient of corporate development of cities in the (a) five years before 2013 and the (b) five years after 2013.

among Hangzhou, Shaoxing, Ningbo, Taizhou, Wenzhou, Lishui, Yiwu, and Quzhou have been closely strengthened. Other major cities have also been further developed. It can be seen that the rapid development of high-speed rail has an important impact on the spatial redistribution of economic activities and strengthens the status of core cities. [21].

#### 6. Conclusions and Future Research Directions

According to the latest research progress in the fields of big data, artificial intelligence, and smart city, the spatial-temporal distribution and registered economic scale of companies in the Yangtze River Delta in recent 40 years are analyzed in this paper. Taking railway station as the center, the changes in the distribution of companies before and after the opening of the railway are analyzed through deep learning, and the relationship of urban development and the opening of the railway were explored through the statistical analysis and correlation analysis. Using ArcGIS engine and C# to visualize big data, the results of spatial clustering of three provinces and one city in the Yangtze River Delta are displayed (Figure 2).

The results in this paper fully illustrate that new technologies such as big data, artificial intelligence, and deep learning have broad prospects in the field of intelligent city applications. Realizing the economic development forecast before and after railway construction can even reveal the inherent law of coordinated development of urban agglomerations and provide data support for urban planning and construction. The use and application of such analysis in this paper can be tested for use and application in the context of other smart cities for specific aspects or scale.

The limitations of the paper are mainly two aspects. One is the in-depth analysis in the text, mostly based on the results of data statistics and calculations, lacking clear data interpretation and effective use by experts in related fields. Another limitation is the prediction results produced by the

#### Complexity

deep learning model. Since this paper is the first application of such data, there is a lack of comparison criteria, and only the effectiveness of the deep learning algorithm can be considered from the perspective of the range of data loss values. Going forward, we may further invite experts in relevant fields to use the data in a more complete and comprehensive way to achieve the ultimate goal of serving smart cities.

#### **Data Availability**

The data used to support the findings of this study are available from the corresponding author upon request.

#### **Conflicts of Interest**

The authors declare no conflicts of interest.

#### Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant nos. 61271351 and 41827807).

#### References

- R. K. R. Kummitha and N. Crutzen, "How do we understand smart cities? An evolutionary perspective," *Cities*, vol. 67, pp. 43–52, 2017.
- [2] K. Su, L. Jie, and H. Fu, "Smart city and the applications," in Proceedings of the 2011 International Conference on Electronics, Communications and Control (ICECC), pp. 1028– 1031, Ningbo, China, September 2011.
- [3] A. Gaur, B. Scotney, G. Parr, and S. Mcclean, "Smart city architecture and its applications based on IoT," *Procedia Computer Science*, vol. 52, pp. 1089–1094, 2015.
- [4] UN, 68% of the World Population Projected to Live in Urban Areas by 2050, Says UN, UN, New York, NY, USA, 2018, https://www.un.org/development/desa/en/news/population/ 2018-revision-of-world-urbanization-prospects.html.
- [5] U. Ekman, "Smart city planning," International Journal of E-Planning Research, vol. 7, no. 3, pp. 1–21, 2018.
- [6] U. Sengupta, C. Doll, A. Gasparatos et al., "Sustainable smart cities: applying complexity science to achieve urban sustainability," UNU-IAS Policy Brief Series, vol. 12, p. 4, 2017.
- [7] S. A. Zhu, "Survey of data-centric smart city research," Information System Engineering, vol. 51, pp. 239–259, 2018.
- [8] J. Wang, C. Li, Z. Xiong, and Z. Shan, "Survey of data-centric smart city," *Journal of Computer Research and Development*, vol. 2, pp. 239–259.
- [9] M. Mohammadi and A. Al-Fuqaha, "Enabling cognitive smart cities using big data and machine learning: approaches and challenges," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 94–101, 2018.
- [10] S. Sagiroglu and D. Sinanc, "Big data: a review," in Proceedings of the 2013 International Conference on Collaboration Technologies and Systems (CTS), pp. 42–47, San Diego, CA, USA, May 2013.
- [11] Z. Xie, X. Ye, Z. Zheng et al., "Modeling polycentric urbanization using multisource big geospatial data," *Remote Sensing*, vol. 11, no. 3, p. 310, 2019.
- [12] B. P. L. Lau, S. H. Marakkalage, Y. Zhou et al., "A survey of data fusion in smart city applications," *Information Fusion*, vol. 52, pp. 357–374, 2019.

- [13] F. Zhao, "Thoughts on the countermeasures of Shanghai surveying and mapping geographic information service for smart cities," *Shanghai Land Resources*, vol. 38, pp. 93–95, 2017.
- [14] D. T. T. Limited, Super Smart City—Happier Society with Higher Quality, Deloitte, New York, NY, USA, https://www2. deloitte.com/cn/en/pages/public-sector/articles/super-smartcity.html.
- [15] Z. Miao-Xi, "Evolution of network and spatial structure in Yangtze river delta," *Geographical Research*, vol. 30, no. 2, pp. 119–131, 2011.
- [16] Q. Shui-Tu and J. F. Jiao, "Structure, industrial agglomeration and regional economic growth: based on Yangtze river delta panel data from 2000 to 2007," *Journal of Business Economics*, vol. 1, pp. 67–74, 2010.
- [17] D. Zhuang, Y. Yang, S. Jin, and R. Han, "Evolution of the Yangtze river delta's city network based on the strategic emerging industries," *Scientia Geographica Sinica*, vol. 37, pp. 546–553, 2017.
- [18] H. Yu, "Strategic conception of integrating Shanghai metropolitan area into the development of Yangtze river delta urban agglomeration," *Shanghai Urban Management*, vol. 2, pp. 20–23, 2018.
- [19] J. Li, "The inherent law of China's economic performance and its future trend of development," *Theory Journal*, vol. 1, pp. 46–54, 2014.
- [20] C. C. Cheng, "The overall framework and strategy of the Yangtze river economic zone's coordinated equilibrium development," *Journal of Nantong University*, vol. 1, pp. 1–8, 2015.
- [21] X. Li, B. Huang, R. Li, and Y. Zhang, "Exploring the impact of high speed railways on the spatial redistribution of economic activities—Yangtze River Delta urban agglomeration as a case study," *Journal of Transport Geography*, vol. 57, pp. 194–206, 2016.
- [22] H. Chang-Fu and X. Yuan, "Research on the role of urban rail transit in promoting economic development," *Procedia En*gineering, vol. 21, pp. 520–525, 2011.
- [23] R. D. Mallon, "Transport and economic development," *Economic Digest*, vol. 3, pp. 8–13, 1960.
- [24] Y. T. Mohmand, A. Wang, and A. Saeed, "The impact of transportation infrastructure on economic growth: empirical evidence from Pakistan," *Transportation Letters*, vol. 9, no. 2, pp. 63–69, 2017.
- [25] P. F. Bangwayo-Skeete and R. W. Skeete, "Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach," *Tourism Management*, vol. 46, pp. 454–464, 2015.
- [26] T. M. V. Kumar and B. Dahiya, Smart Economy in Smart Cities, Springer, Singapore, 2017.
- [27] J. Chan, "The effect of college education on intolerance: evidence from google search data," *Applied Economics Letters*, vol. 26, no. 2, pp. 83–86, 2018.
- [28] R. Wujun and L. I. Xin, "Tourism demand analysis based on Internet big data: the case of Huairou, Beijing," Systems Engineering-Theory & Practice, vol. 38, no. 2, pp. 437–443, 2018.
- [29] H.-W. Wang and D.-R. Chen, "Economic recession and obesity-related Internet search behavior in taiwan: analysis of google trends data," *JMIR Public Health and Surveillance*, vol. 4, no. 2, p. e37, 2018.
- [30] M. Scriney, M. F. O'Connor, and M. Roantree, "Generating cubes from smart city web data," in *Proceedings of the Australasian Computer Science Week Multiconference*, pp. 1–8, Geelong, Australia, January 2017.

- [31] R. Sen, D. Quercia, and K. P. Gummadi, "Scalable urban data collection from the web," in *Proceedings of the AAAI*, San Francisco, CA, USA, February 2017.
- [32] B. J. Alberto, L. J. Emilio, Q. Enedina et al., "A web-based tool for automatic data collection, curation, and visualization of complex healthcare survey studies including social network analysis," *Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 2579848, 8 pages, 2017.
- [33] A. Afaneh and I. Shahrour, "Use of GIS for SunRise smart city project, large scale demonstrator of the smart city," in *Proceedings of the 2017 Sensors Networks Smart and Emerging Technologies (SENSET)*, pp. 1–4, Beirut, Lebanon, September 2017.
- [34] S. Yamamura, L. Fan, and Y. Suzuki, "Assessment of urban energy performance through integration of BIM and GIS for smart city planning," *Procedia Engineering*, vol. 180, pp. 1462–1472, 2017.
- [35] M. Maloof, "Some basic concept of machine learning and data mining," in *Machine Learning and Data Mining for Computer Security*, Springer, London, UK, 2006.
- [36] K. J. Bergen, P. A. Johnson, M. V. de Hoop, G. C. Beroza, and G. C. Beroza, "Machine learning for data-driven discovery in solid Earth geoscience," *Science*, vol. 363, no. 6433, p. eaau0323, 2019.
- [37] M. I. Jordan and T. M. Mitchell, "Machine learning: trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.
- [38] Y. Tarutani, K. Hashimoto, G. Hasegawa et al., "Temperature distribution prediction in data centers for decreasing power consumption by machine learning," in *Proceedings of the IEEE International Conference on Cloud Computing Technology & Science*, pp. 635–642, Vancouver, Canada, November 2015.
- [39] J. Chin, V. Callaghan, and I. Lam, "Understanding and personalising smart city services using machine learning, the internet-of-things and big data," in *Proceedings of the 2017 IEEE 26th International Symposium on Industrial Electronics* (ISIE), pp. 2050–2055, Edinburgh, UK, June 2017.
- [40] N. Lei, Z. Luo, S. T. Yau, and D. X. Gu, "Geometric understanding of deep learning," 2018, https://arxiv.org/abs/ 1805.10451.
- [41] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [42] Y. Kai, J. Lei, C. Yuqiang, X. Wei, and Baidu, "Deep learning: yesterday, today, and tomorrow," *Journal of Computer Re*search and Development, vol. 20, no. 6, p. 1349, 2013.
- [43] C. Balchandani, R. K. Hatwar, P. Makkar, Y. Shah, and M. Eirinaki, "A deep learning framework for smart street cleaning," in *Proceedings of the 2017 IEEE Third International Conference on Big Data Computing Service and Applications* (*BigDataService*), pp. 112–117, San Francisco, CA, USA, April 2017.
- [44] C. Dongjie, Z. Wensheng, and Y. Yang, "Detection and recognition of high-speed railway catenary locator based on deep learning," *Journal of University of Science and Technology* of China, vol. 47, pp. 320–327, 2017.
- [45] A. Camero, T. Jamal, D. H. Stolfi, and E. Alba, "Evolutionary deep learning for car park occupancy prediction in smart cities," in *Proceedings of the 12th International Conference, LION*, Kalamata, Greece, June 2019.
- [46] I. A. T. Hashem, V. Chang, N. B. Anuar et al., "The role of big data in smart city," *International Journal of Information Management*, vol. 36, no. 5, pp. 748–758, 2016.
- [47] S. Shams, S. Goswami, K. Lee, S. Yang, and S. J. Park, "Towards distributed cyberinfrastructure for smart cities using

big data and deep learning technologies," in *Proceedings of the* 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS), pp. 1276–1283, Vienna, Austria, July 2018.

- [48] D. Myers and J. W. McGuffee, "Choosing scrapy," *Journal of Computing Sciences in Colleges*, vol. 31, pp. 83–89, 2015.
- [49] C. Stephanidis, G. Salvendy, M. Antona et al., "Seven HCI grand challenges," *International Journal of Human-Computer Interaction*, vol. 35, no. 14, pp. 1229–1269, 2019.
- [50] A. B. A. P. Indyk, "Which regular expression patterns are hard to match?," in *Proceedings of the 2016 IEEE 57th annual* symposium on foundations of computer science (FOCS), pp. 457–466, New Brunswick, NJ, USA, October 2016.
- [51] Z. Guojun, J. W., S. Jihui, S. Fan, Z. Hao, and L. Jiang, "Design and application of intelligent dynamic crawler for web data mining," in *Proceedings of the 2017 32nd Youth Academic Annual Conference of Chinese Association of Automation* (YAC), pp. 1098–1105, Hefei, China, May 2017.
- [52] F. Gessert, W. Wingerath, S. Friedrich, and N. Ritter, "NoSQL database systems: a survey and decision guidance," *Computer Science—Research and Development*, vol. 32, no. 3-4, pp. 353–365, 2017.
- [53] H. Jing, E. Haihong, L. Guan, and D. Jian, "Survey on NoSQL database," in Proceedings of the International Conference on Pervasive Computing & Applications, pp. 363–366, Port Elizabeth, South Africa, October 2011.
- [54] J. Zhu, J. Wang, and J. Zhang, "Research and application of big data query technology based on NoSQL database," *China Railway Science*, vol. 35, no. 1, pp. 135–141, 2014.
- [55] Y. S. Kang, I. H. Park, J. Rhee, and Y. H. Lee, "MongoDBbased repository design for IoT-generated RFID/sensor big data," *IEEE Sensors Journal*, vol. 16, no. 2, pp. 485–497, 2015.
- [56] S. Li, H. Yang, Y. Huang, and Q. Zhou, "Geo-spatial big data storage based on NoSQL database," *Geomatics and Information Science of Wuhan University*, vol. 42, pp. 163–169, 2017.
- [57] M. Panagiotopoulou and A. Stratigea, "Spatial data management and visualization tools and technologies for enhancing participatory E-planning in smart cities," in *Smart Cities in the Mediterranean*, Springer, Cham, Switzerland, 2017.
- [58] J. A. Hartigan and M. A. Wong, "Algorithm as 136: a k-means clustering algorithm," *Applied Statistics*, vol. 28, no. 1, pp. 100–108, 1979.
- [59] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Transactions* on Pattern Analysis and Machine Intelligence, vol. 24, pp. 0– 892, 2002.
- [60] A. Radhakrishna, S. Appu, S. Kevin, L. Aurelien, F. Pascal, and S. Sabine, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 34, pp. 2274–2282, 2012.
- [61] X. Lin, H. Q., H. Zhang, and Y. Huang, "The 5I model of smart city: a case of Shanghai, china," in *Proceedings of the 2015 IEEE First International Conference on Big Data Computing Service and Applications*, pp. 329–332, Redwood City, CA, USA, March 2015.
- [62] D. M. Hawkins, "The problem of overfitting," *Journal of Chemical Information and Computer Sciences*, vol. 44, no. 44, pp. 1–12.
- [63] Z. E. Ross, M.-A. Meier, and E. Hauksson, "PWave arrival picking and first-motion polarity determination with deep

learning," Journal of Geophysical Research: Solid Earth, vol. 123, no. 6, pp. 5120-5129, 2018.

- [64] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," in *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 2672– 2680, Montreal, Canada, December 2014.
- [65] P. Mainali and R. Wittebrood, Machine Learning Deconvolution Filter Kernels for Image Restoration, SPIE, Bellingham, WA, USA, 2015.
- [66] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, https://arxiv.org/abs/1512. 03385.
- [67] J. Li, Y. Huang, and Y. Gong, "Improved cepstra minimummean-square-error noise reduction algorithm for robust speech recognition," in *Proceedings of the IEEE International Conference on Acoustics*, pp. 4865–4869, New Orleans, LA, USA, March 2017.
- [68] T. Pang, C. Du, and J. Zhu, "Robust deep learning via reverse cross-entropy training and thresholding test," 2017, https:// arxiv.org/abs/1706.00633.
- [69] H. V. Nguyen and L. Bai, "Cosine similarity metric learning for face verification," in *Proceedings of the Asian Conference* on Computer Vision, pp. 709–720, Queenstown, New Zealand, November 2010.
- [70] S. Kala, S. Nalesh, B. R. Jose, and J. Mathew, "Image reconstruction using novel two-dimensional fourier transform," in Advances in Soft Computing and Machine Learning in Image Processing, A. E. Hassanien and D. A. Oliva, Eds., pp. 699–718, Springer International Publishing, Cham, Switzerland, 2018.
- [71] D. Wang, H. Lu, and C. Bo, "Visual tracking via weighted local cosine similarity," *IEEE Transactions on Cybernetics*, vol. 45, no. 45, pp. 1838–1850, 2015.
- [72] Z. Tao, H. Liu, H. Fu, and Y. Fu, "Image cosegmentation via saliency-guided constrained clustering with cosine similarity," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, San Francisco, CA, USA, February 2017.
- [73] R. Siedlecki and D. Papla, "Conditional correlation coefficient as a tool for analysis of contagion in financial markets and real economy indexes based on the synthetic ratio," *Procedia—Social and Behavioral Sciences*, vol. 220, pp. 452–461, 2016.
- [74] Z. Haoqiang, F. Shanshan, Z. Di, and M. Zhe, "Using Bayesian network for risk analysis of inland waterway transportation system considering environmental effects," in *Proceedings of* the International Conference on Transportation Information & Safety, pp. 689–694, Wuhan, China, June 2015.
- [75] Z. Guang-Yuan, D. Xiao-Dong, L. U. Gong-Yuan, and Y. U. Yang, "Anexperimental analysison the turnover time of railway freight transportation based on the impact strength of factors," *Journal of Chongqing University (English Edition)*, vol. 16, pp. 113–123, 2017.
- [76] J. S. Liu, C. Li, S. T. Li, H. M. Hong, and K. Hu, "Predicting the smart meters life cycle based on the analysis of correlation coefficient," *Applied Mechanics and Materials*, vol. 313-314, pp. 629-633, 2013.
- [77] Y. Hu, "The Impact of the sixth train speed increase on the Yangtze River Delta," *Yangtze River Delta*, vol. 5, pp. 12-13, 2007.
- [78] J. Zong, T. Liu, Q. Gong, and B. Wang, "Dynamic response study on speed-increase subgrade of existing railway," *China Railway Science*, vol. 28, pp. 7–11, 2007.
- [79] China, n. b. o. s. o, http://data.stats.gov.cn/ks.htm? cn=E0103&zb=A0101&reg=3100002019.



**Operations Research** 

International Journal of Mathematics and Mathematical Sciences







Applied Mathematics

Hindawi

Submit your manuscripts at www.hindawi.com



The Scientific World Journal



Journal of Probability and Statistics







International Journal of Engineering Mathematics

Complex Analysis

International Journal of Stochastic Analysis



Advances in Numerical Analysis



**Mathematics** 



Mathematical Problems in Engineering



Journal of **Function Spaces** 



International Journal of **Differential Equations** 



Abstract and Applied Analysis



Discrete Dynamics in Nature and Society



Advances in Mathematical Physics