Analysis of Radar Technology Identification Model for Potential Geologic Hazard based on Convolutional Neural Network and Big Data

Feng He
Yunan University of Finance and Economics
Hongjiang Liu (zz1923@ynufe.edu.cn)
Leshan Normal University
Chunxue Liu
Yunan University of Finance and Economics
Guangjing Bao
Yunan University of Finance and Economics

Research Article

Keywords: big data analysis, CNN, Apriori algorithm, geological hazard, radar identification technique

Posted Date: April 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-371548/v1

Analysis of Radar Technology Identification Model for Potential Geologic Hazard based on Convolutional Neural Network and Big Data

Feng He [1], Hongjiang Liu [2]*, Chunxue Liu[1], Guangjing Bao[1]

[1] School of Urban and Environmental Sciences, Yunnan University of Finance and Economics, Kunming, 650221, China;

[2] School of Tourism, Leshan Normal University, Leshan, 614000, China

*Correspondence: zz1923@ynufe.edu.cn

Abstract: To ensure the proper adoption of new technologies in identifying the potential geologic hazard on tourist routes, convolutional neural network (CNN) technology is applied in the radar image geologic hazard information extraction. A scientific and practical geologic hazard radar identification model is built, which is based on CNN's image identification and big data algorithm calculation, and it can effectively improve the geologic hazard identification accuracy. By designing experiments, the geologic hazard radar image data are verified, and the practicality of radar image intelligent Identification under CNN and big data technology is also verified. The results show that the images of different resolution sizes all play a significant role in identification of geologic hazard performed by CNN. However, there are differences in the performance of different CNN models. With the continuous increase of training samples, the identification accuracy of various network models is also improved. By means of radar image test, the identification capability of CNN model is the best, the highest precision is 93.61%, and the geologic hazard recall rate is 98.27%. Apriori algorithm is introduced into data processing, and the running speed and efficiency of identification models are improved, with favorable identification effect in variable data sets. This research can provide theoretical ideas and practical value for the development of potential geologic hazard identification on tourist routes.

Keywords: big data analysis; CNN; Apriori algorithm; geological hazard; radar identification technique

1. Introduction

With the development of economic society, tourism has gradually become a hot industry. It is a comprehensive and diversified industrial collection, which plays an important role in economic and social development [1]. Many tourist sites are located in the mountains, rocks, and other complex terrain, which often leads to the occurrence of geologic hazard [2]. The geologic hazard is an event that occurs on the surface of the earth and mainly depends on the movement of rocks [3]. The occurrence of geologic hazard is often due to the destruction of nature by humans, which breaks the original ecological balance. It not only affects the safety of life and property, but also causes irreversible damage to the human living environment [4]. Common geologic hazards include collapse, landslide, mudslide, ground subsidence, ground fissure, and ground subsidence, which all pose a serious threat to life and property [5]. Geologic hazard will not only bring immeasurable losses to a country or individual's economy, it will also greatly destroy people's living conditions and stability. Therefore, the prevention of geologic hazard has become an issue of increasing concern [6]. The most advanced technology should be adopted to establish corresponding technical measures, and geologic hazard emergencies should be prevented efficiently and timely, so that people can avoid the losses caused by geologic hazard [7]. For the prevention and treatment of geologic hazard, it doesn't just refer to simple preventive prediction or engineering governance, and how to use radar data to analyze potential geologic hazard is an important protective measure for protecting people's life and property [8]. Therefore, it is very meaningful to adopt big data and deep learning technology to build a geologic hazard model.

With the continuous development of deep learning technology in recent years, more and more practical problems have been solved. In deep learning, there is a CNN that is widely used in model building, especially in the field of image identification [9]. At first, it was only used in natural image identification, which got good identification results [10]. With the continuous progress of radar technology, the spatial resolution of radar images is getting closer and closer to natural pictures. Through geometric correction, radiometric calibration, and red, green, and blue (RGB) band image combination, the images are also similar to natural pictures in terms of deformation and color. Therefore, CNN technology has also achieved good results in interpretation of radar image [11]. Zhou et al. (2016) [12] developed many convolutional neural radar image classifiers, such as Softmax regression model and

support vector machine. Shen et al. (2019) [13] adopted the multi-level semantics extracted via CNN, which had strong generalization ability and good expression effect, and was significantly better than other hand-designed features. Can et al. (2019) [14] adopted Visual Geometry Group Network 16 and mobile Web-based Geographic Information Science as the basic model, and high performance was realized by choosing the best hyperparameters, with the highest precision of 94%. The establishment of the potential geologic hazard predication model is the most critical step for the smooth application of CNN technology, and its quality will directly affect the final identification effect [15]. It requires a large number of samples as a data basis, but for present, there are relatively few studies in this field [16]. Therefore, conducting research in this field is of important theoretical significance and practical application value.

In geologic hazard monitoring, it is of great significance to find the changes of various geological conditions and other factors in various geographical environments in time, and make efficient, fast, and safe judgments. Then, preventive measures can be conducted in time to avoid geologic hazard accidents and protect the safety of people's lives and properties. The historical data and monitoring information obtained from the geologic hazard monitoring database can reasonably and accurately analyze the reasons for the change. Moreover, high-quality, high-level research theoretical methods should be applied for exploring potentially valuable spatial knowledge from multiple visions, which can provide decision support for China's Geological Bureau and various relevant departments.

2. Materials and Methods

2.1 Traditional geologic hazard model

The geologic hazard model was originally based on statistical modeling of data collected from man-made trips or weather stations. The system includes sensors, transmission, monitoring, calculation, and display modules, among which sensors can obtain real-time geological environmental data. Transmission module is for transmitting the acquired geological environment data to the monitoring module through the network. Monitoring module is for receiving the geological environment data output by the transmission module, and converting the environmental data into the identification format of the calculation module, which are then output to the calculation module. Calculation module is for calculating the environmental data and obtaining geological environmental results, which output the geological environment result to the display module. The display module outputs the geological environment results, and outputs two-dimensional and three-dimensional result maps that characterize the geological environment based on the geological environment results [17]. The specific process is shown in Figure 1.

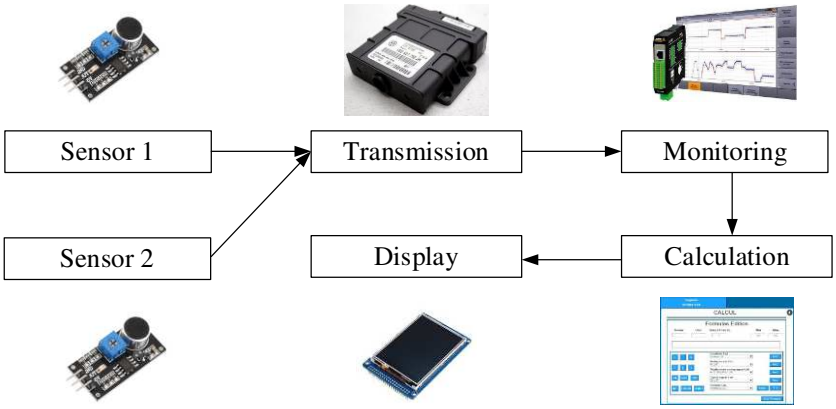

Figure. 1 Traditional geologic hazard identification model.

The traditional geologic hazard identification model is based on manual field investigation and survey. Although this manual operation method is easier to obtain the actual and accurate data information of the survey area, yet it requires a lot of manpower and material resources, and the real-time performance is low. Moreover, the life of investigators can't be guaranteed, and it is quite dangerous. Manual surveys can't be carried out smoothly in some areas where traffic is severely blocked and difficult to reach [18]. In this dilemma, radar detection technology, which is with the characteristics of fast data information acquisition, short period, easy operation, massive information, and less restricted conditions, has become a new survey approach [19]. The image enhancement can effectively improve the definition of the image, which is beneficial to obtain the useful data in the image. It can effectively improve the contrast of captured images at night and reduce interference noise in the image. Moreover, it can

effectively solve the image degradation and fogging phenomenon caused by the interference of atmospheric scattering under severe weather such as fog and heavy rain [20].

2.2 CNN

CNN is a feedforward neural network, whose artificial neuron can respond to a part of the surrounding cells in the coverage area. It is composed of one or more convolutional layers and a fully connected layer at the top, which also includes the associated weight and pooling layer. In Figure 2, this structure enables the CNN to adopt the two-dimensional structure of the input data. Compared with other deep learning structures, CNN can get better results in image identification, which can also be trained by back-propagation algorithms. CNN is mainly divided into input layer, output layer, and hidden layer. The hidden layer includes convolutional layer, pooling layer, and fully connected layer, among which the hidden layer is the most important.

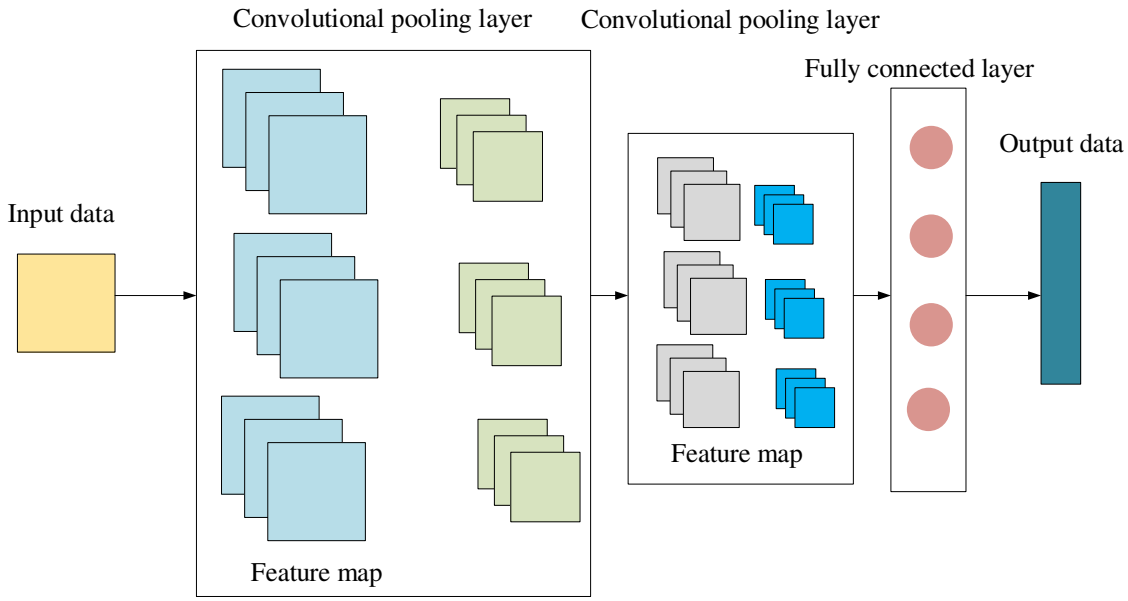

Figure. 2 Schematic diagram of CNN structure.

The main function of input layer of CNN is preprocessing the original radar image data (Figure 3), which includes de-averaging and normalization. De-averaging: for features of each training image, the feature average of all training data is subtracted, then the centers of all training samples are at the origin of coordinates. Normalization: after processing, the features of all dimensions of the data are in the same range of change, and the deviation caused by the different value ranges is reduced. For example, the range of feature A is 0 to 100, and the range of feature B is 0 to 100,000, it is wrong to use them directly without any treatment. At this time, it is necessary to normalize the data of A and B to the range of 0 to 1. PCA/whitening: PCA dimensionality reduction means reducing the correlation between input data. Whitening is similar to normalization, that is, processing the variation range of data features in each dimension into the same interval. The structure and working mechanism of the CNN output layer are the same as the output layer in the traditional feedforward neural network. For image classification, the output layer adopts a logistic function or a normalized exponential function to output category labels. In the object identification, the output layer can output the center coordinates, size, and category of the target object. In image semantic segmentation, the output layer directly outputs the classification result of each pixel.

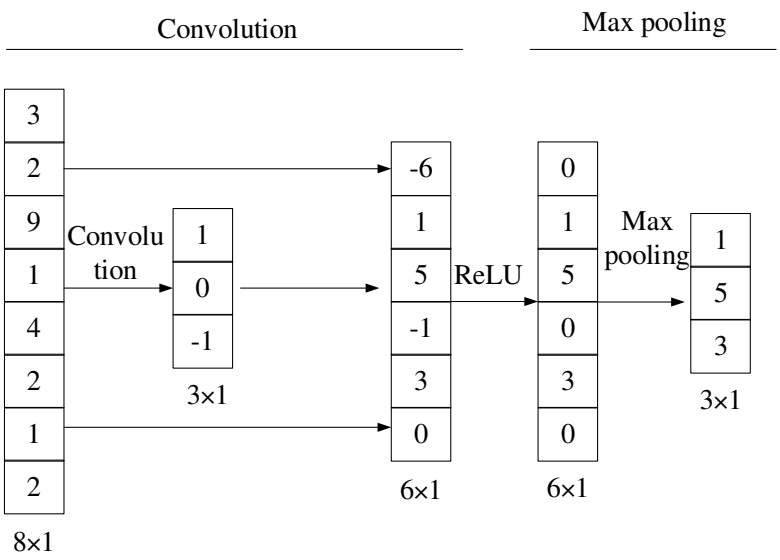

Figure. 3 Schematic diagram of data pre-processing in CNN input layer.

The convolutional layer extracts features from the input data, and the convolutional layer is the most important part of the CNN. Compared with ordinary neural networks, the input layer and hidden layer are designed as "fully connected layers". The convolutional layer restricts the connections between neurons and input data, that is, each neuron only connects a small part of the input image. Since the neurons of the convolutional layer are also three-dimensional, they also have depth. The convolutional layer parameters include a series of filters, each of which trains one depth, and there are several filter output units that have as much depth. The input unit size is 128×128×3, and the output unit depth is 5. For the same site with different depths of the output unit, the area connected to the input picture is the same, but the parameter filter is different. The calculation of the convolutional layer is as follows.

$$C_i = f(C_{i-1} * W^i + b^i) \tag{1}$$

In equation (1), C_i represents the feature map of the *ith* layer, W^i and b^i represent the weight and bias vector of the *ith* layer, respectively. * represent the convolution operation of the convolution kernel and the input image, and *f* represents the activation function. The commonly used activation function is ReLu function, and its expression is as follows.

$$f(x) = \begin{cases} x, & x > 0 \\ 0, & x \le 0 \end{cases} \tag{2}$$

The pooling layer is located between the convolutional layers, and its main function is compressing the input data and parameters to prevent overfitting. In short, if the input is an image, then the main function of the pooling layer is compressing the image, as shown in Figure 4. The calculation of the pooling layer is as follows.

$$Hi = p(H_{i-1}) \tag{3}$$

In equation (3), *Hi* represents the output of the *ith* layer, and *p* represents the pooling function. The purpose of the fully connected layer is mapping the features learned through multiple convolutions and pooling to the label space of the sample, whose function is the same as the hidden layer in the traditional feedforward neural network that all neurons between the two layers have weight connection. It can convert the two-dimensional feature map output by the convolution into a one-dimensional vector, and pass it to the output layer through the excitation function for final classification.

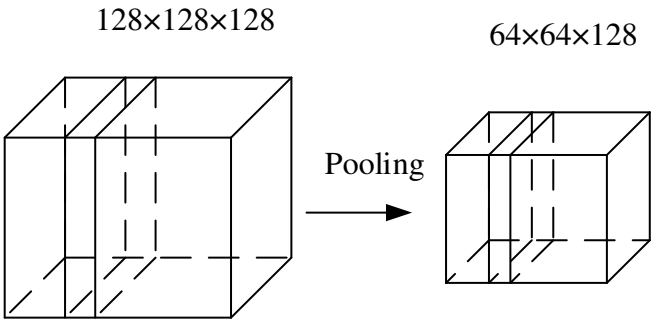

Figure. 4 Schematic diagram of CNN pooling layer.

To verify the advantages and disadvantages of the algorithm, the following neural networks are adopted for comparison. I: LeNet-5, it uses convolution, parameter sharing, pooling, and other operations to extract features, avoiding a lot of computational costs, and it finally uses a fully connected neural network for classification and identification [21]. II: AlexNet, it has 8 layers, including 5 layers of convolutional layers and 3 layers of fully connected layers, and finally the softmax layer with 1000 neurons receives its output, corresponding to the distribution of 1,000 category labels [22]. III: VGGNet, it successfully builds a 16-19 layers CNN by repeatedly stacking a convolution kernel with the scale of 3×3 and the maximum pooling layer with the scale of 2×2. It is proved that increasing the depth of the network can improve the final performance of the network to a certain extent, which also greatly reduces the error rate, and improves the generalization ability of the network [23]. IV: GoogLeNet, it adopts the concept of "Inception" module proposed in "Network in network", which has 22 layers in depth, but its size is much smaller than AlexNet and VGG. GoogleNet has 5 million parameters, and when memory or computing resources are limited, GoogLeNet is a better choice [24]. V: ResNet, the idea of ResNet is very similar to that of Highway Network, which allows the original input information to be directly transmitted to the subsequent layers. The neural network of this layer can learn the residual of the previous network output instead of learning the entire output [25].

2.3 Big data calculation methods

It is also crucial for the processing of geologic hazard data, and Apriori algorithm is chosen to be adopted. The algorithm first scans the transaction data sets when mining them. It counts each item set and compare it with the minimum support count set by the user. If the number of item sets is greater than the minimum support count, then the item sets are regarded as frequent item sets; but if it is smaller, it will be deleted. Then, the frequent item sets obtained from the first scan are self-connected to obtain candidate item sets, the transaction data set is scanned again, and the candidate item sets obtained by the connection are counted. If it is larger than the minimum support count, then the item set is regarded as a frequent item set, but if it is smaller, it will be deleted. This step will be repeated continuously until no $k+1$ candidate item set is generated, and the final kth item set is the final frequent k items set [26].

The algorithm has two key steps, namely the connection step and the pruning step. The connection step finds k frequent item sets and performs self-connection operations on them to generate new candidate frequent item sets. The pruning step scans the database first when mining frequent item sets, to determine that the count of each candidate in the candidate set is greater than or equal to the minimum support count. Then, frequent item sets are obtained. Finally, the frequent item set is determined according to the count of each candidate in the candidate set [27], and Figure 5 is the algorithm's flow chart.

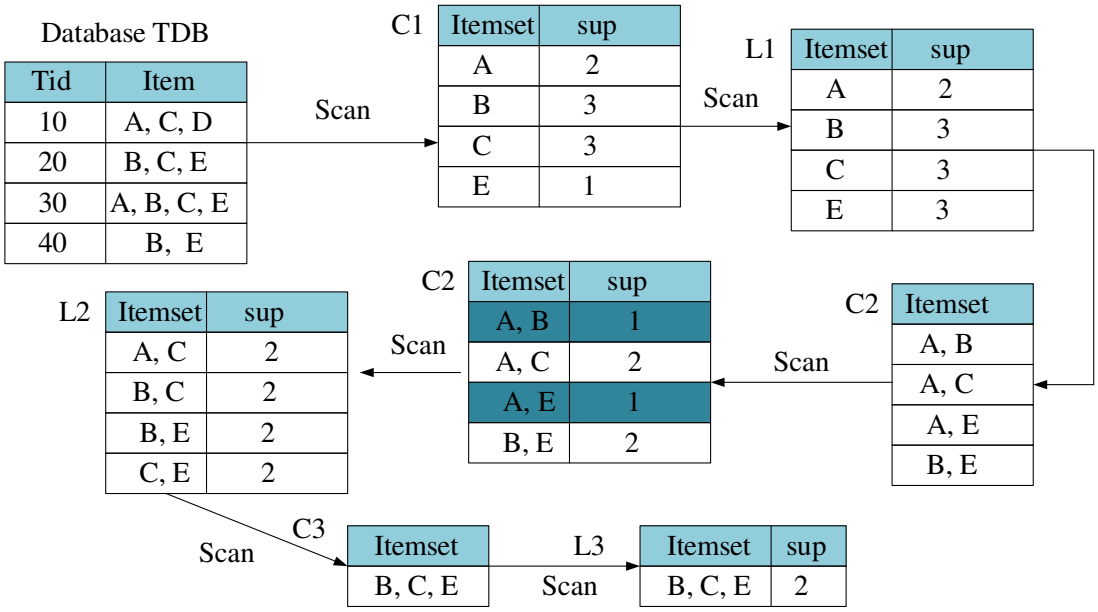

Figure. 5 Flow of Apriori algorithm.

2.4 Data sources and training

I: Data source. The radar data come from two remote sensing satellites Landsat 8 and Sentinel 1. The specific information is shown in Table 1. ENVI5.1 radar image processing software is adopted to preprocess the data of Data 1 and Data 2, as shown in Figure 6, the image resolution is 128×128. To enhance the model's ability to identify disasters at different angles and directions, the image is rotated to different directions, thus creating a larger data set.

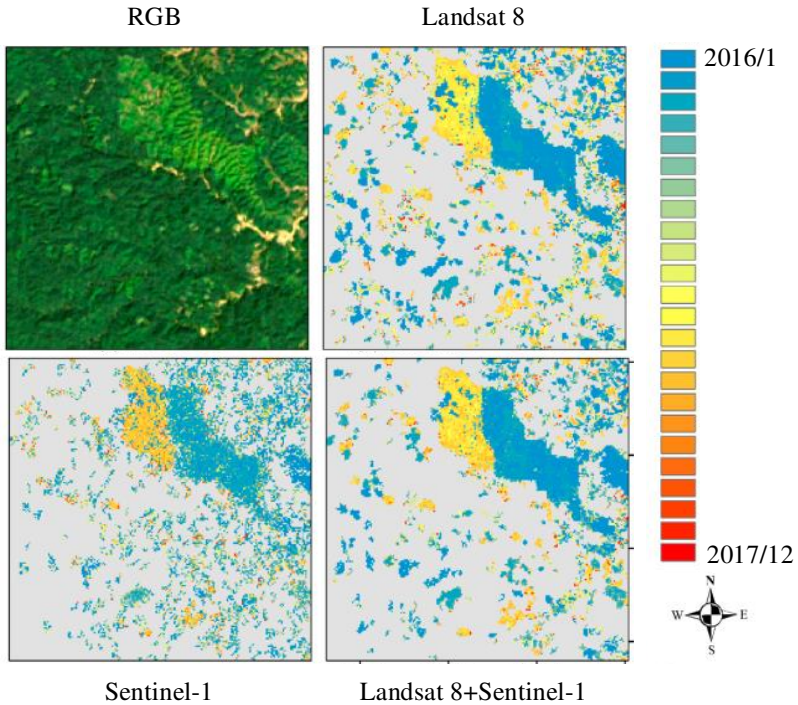

Figure. 6 Sample radar data of Landsat 8 and Sentinel-1.

Table. 1 Data source information of radar image.

Category	Landsat 8 OIL_TIRS+ Sentinel-1	
	Data 1	Data 2
Data identification	LC81220362016246LGN00	LC81210362016207LGN00
Strip number	122/36	121/36

Date	2016-09-02	2016-07-25
Average cloud cover	0.08	2.98
Center latitude and longitude	116.55271, 34.61072	118.09952, 34.61094

II: model training. Since the geologic hazard scene on the tourist route is very complex and the scale size distribution is unbalanced, and the CNN requires the input image size to be consistent, the size of the input sample resolution may have a certain impact on the final identification effect. Therefore, from the perspective of sample resolution, CNN can play the greatest role in the interpretation of geologic hazard radar. Giving that the scale problem of geologic hazard and the storage mechanism of the computer itself, the original samples are divided into five resolutions: 10×10, 50×50, 128×128, 256×256, and 512×512 according to the number of convolutions. According to the ratio of 7:3, the resampled samples are randomly divided into training set and validation set, as shown in Table 2. "0" represents non-geologic hazard, and "1" represents geologic hazard.

Table. 2 Allocation table of training set and validation set.

Training set		Validation set	
0	1	0	1
694	227	298	97

2.5 Evaluation of model performance

For application scenarios such as geologic hazard radar interpretation, it is hoped that the overall classification precision is high, and all geologic hazards should be identified as much as possible. Based on this, the four evaluation indicators of precision, recall rate, accuracy, and comprehensive evaluation are selected to evaluate the performance of the geologic hazard radar identification model [28], and the data performance evaluation indicators are operating efficiency and speedup.

Table. 2 Classification of confusion matrix.

Category	Surface normal class	Surface negative	Total
Predicted positive	TP	FP	TP+FP
Predicted negative	FN	TN	FN+TN
Total	TP+FN	FP+TN	TP+FP+FN+TN

TP represents the number of positive samples predicted to be positive; *FP* represents the number of negative samples predicted to be positive; *FN* represents the number of positive samples predicted to be negative; and *TN* represents the number of negative samples predicted to be negative. Accuracy (ACC) is taken to measure the overall classification accuracy, that is, the proportion of samples that predict correctly, as shown in equation (4).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

Recall rate (Rec): is taken to measure the coverage of positive samples, that is, the proportion of positive samples that are correctly classified to the total positive samples, as shown in equation (5).

$$\mathrm{Rec} = \frac{TP}{TP + FN} \tag{5}$$

The precision rate (Pre) represents the proportion of examples classified as positive examples that are actually positive examples.

$$\mathrm{Pr}e = \frac{TP}{TP + FP} \tag{6}$$

Pre and Recall sometimes have conflicts, so they need to be considered comprehensively. The most common method is F-Measure, which is the weighted harmonic average of precision and recall. The calculation is shown in equation (7).

$$F\text{-}measure = \frac{2 * \mathrm{Pr}e * \mathrm{Re}c}{\mathrm{Pr}e + \mathrm{Re}c} \tag{7}$$

Speedup is the ratio of the time consumed by the same task in a single processor system and a parallel processor system, which is taken to measure the performance and effect of parallel system or program parallelization. The calculation is shown in equation (8).

$$Sp = \frac{T_1}{T_p} \tag{8}$$

Operational efficiency: when the speed-up ratio of a parallel algorithm is the ideal speed-up ratio, if the number of processors is doubled, the execution speed will be doubled. The specific calculation is shown in equation (9).

$$E_p = \frac{S_p}{p} = \frac{T_1}{p \cdot T_p} \tag{9}$$

P is the number of Graphics Processing Unit, T_1 is the time to execute the algorithm, and T_p is the time to execute the algorithm when there are p processors.

3. Results and Discussion

3.1 The influence of data resolution on identification effect

From Figure 7, under the same CNN model, different sample sizes will have certain impact on the precision, recall rate, accuracy, and comprehensive evaluation of the final model. The overall value of the model is distributed between 77.29%-93.61 %, which is at a relatively high level. Among them, the best sample size of the AlexNet, VGGNet, GoogLeNet, ResNet, and CNN is 128×128, 256×256, 128×128, 256×256, and 256×256, respectively. Too large samples will be accompanied by more redundant information, which has a negative impact on the improvement of identification accuracy. The AlexNet network is least affected by the sample block size and the overall accuracy is generally high, while the ResNet network is most affected by the sample block size and the overall accuracy is generally low. Among all CNNs, the CNN network has the best performance, which is maintained at about 90% on average, with the highest accuracy of 93.61%. Based on the above results, the CNN hazard identification model has the best performance when the sample size is 256×256.

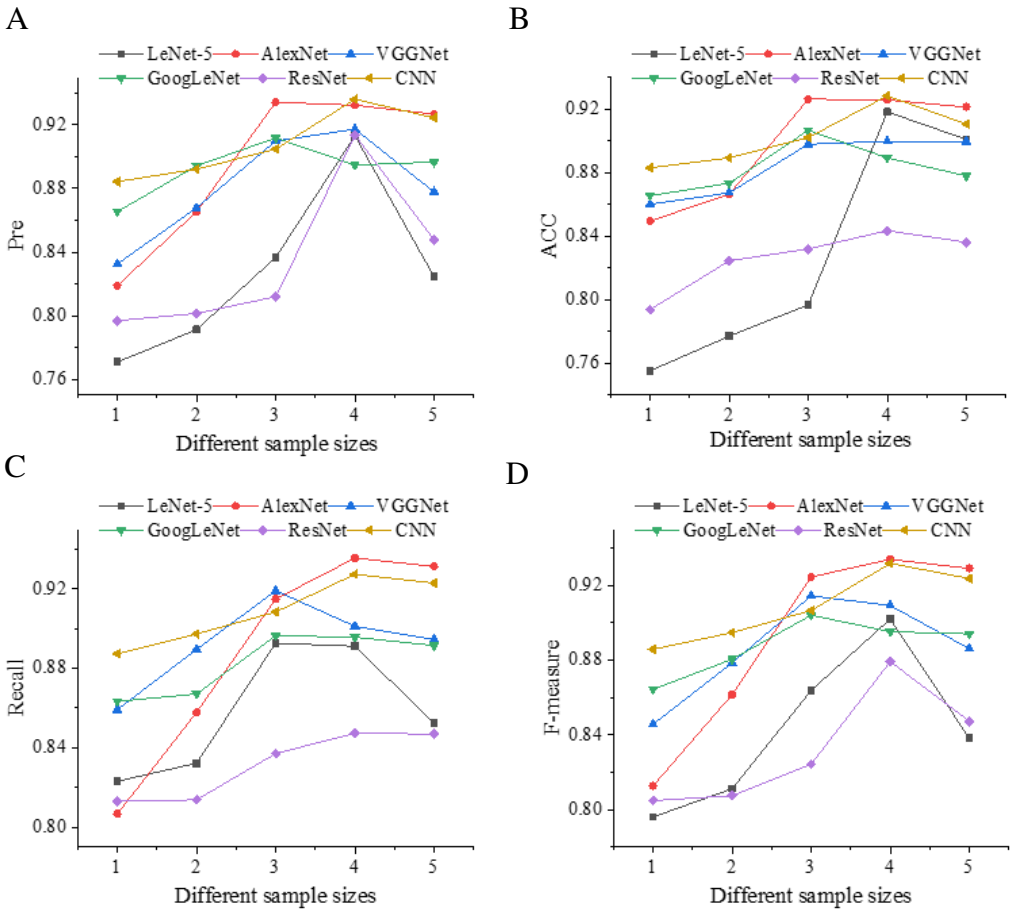

Figure. 7 The effect of sample size on identification effect.
(Note: 1-5 are five resolution sizes of 10×10, 50×50, 128×128, 256×256, and 512×512, respectively.)

3.2 The effect of data size on identification effect

To further verify the impact of data size on performance of model, a certain number of training samples are randomly selected from each of the non-geologic hazards and geologic hazards in the training set according to the ratio of 3:1, and the number interval is set to 100, so as to observe the trend of identification accuracy. From Figure 8, with the increase in the number of training samples, the overall change trend of the network model's precision, recall rate, accuracy, and comprehensive evaluation is constantly increasing. With the increase in the number of training samples, the correlation between samples is also continuously strengthened, and the identification accuracy of the network will stabilize. There may also be interference information, resulting in a decrease in identification accuracy. Among all the models, the CNN network model has the best performance, with the highest accuracy of 92.65%. The above results fully illustrate the validity of the potential geologic hazard radar identification model for the tourist routes.

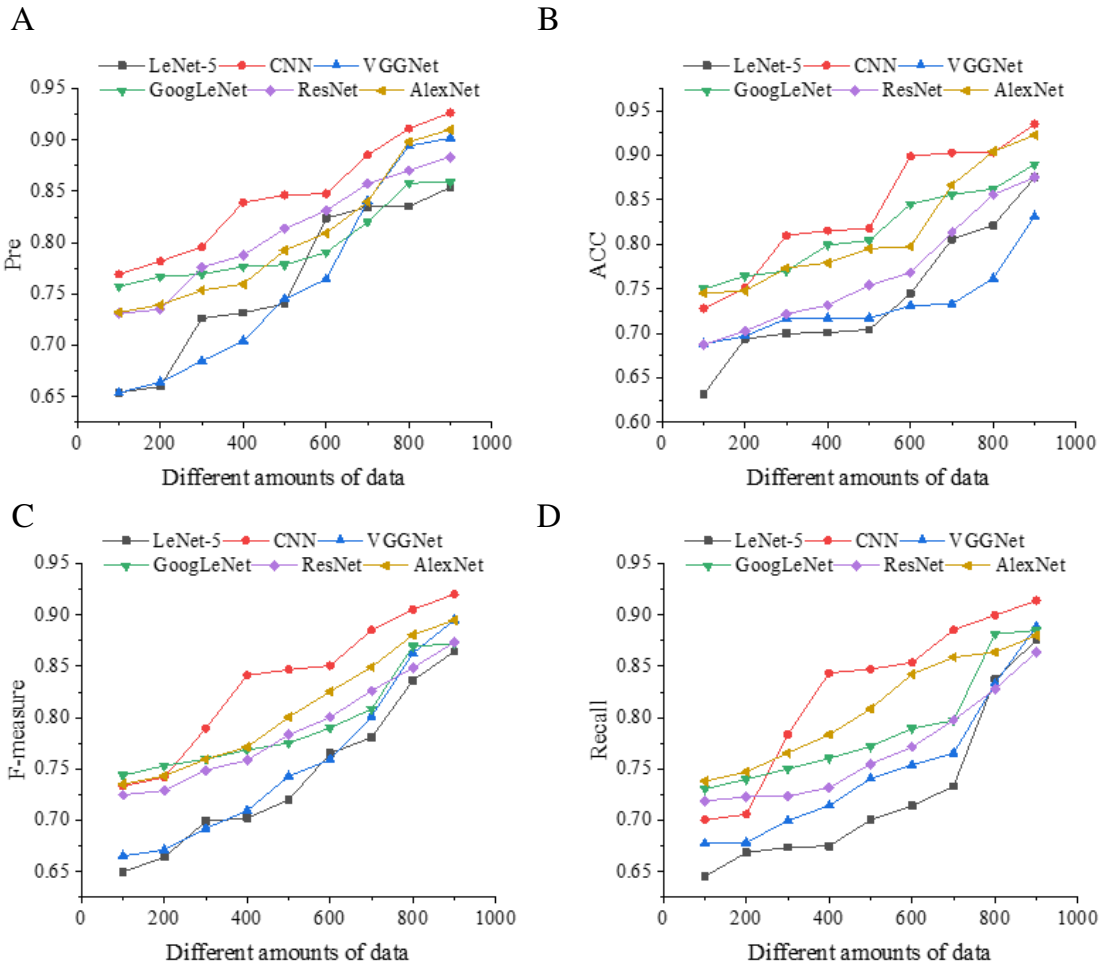

Figure. 8 The effect of data size on identification effect.

3.3 Data processing performance analysis of geologic hazard model

In Figure 9-A, as the number of nodes continues to increase, the model speedup ratio is also increasing. When the data nodes increase to a certain extent, the speedup ratio of the algorithm is significantly lower than that of the data node. The reason is that when the two data sizes are basically the same, the communication overhead between nodes will increase due to the increase of nodes. As a result, the time taken up will increase. In Figure 9-B, three test data sets of different scales are adopted for experiments. The algorithm will increase with the number of data nodes, and the running time of mining association rules and analyzing data will be faster and faster. Under the same scale of test data, as the number of data nodes increases, the operation efficiency of a relatively large scale is obviously higher than that of a small scale.

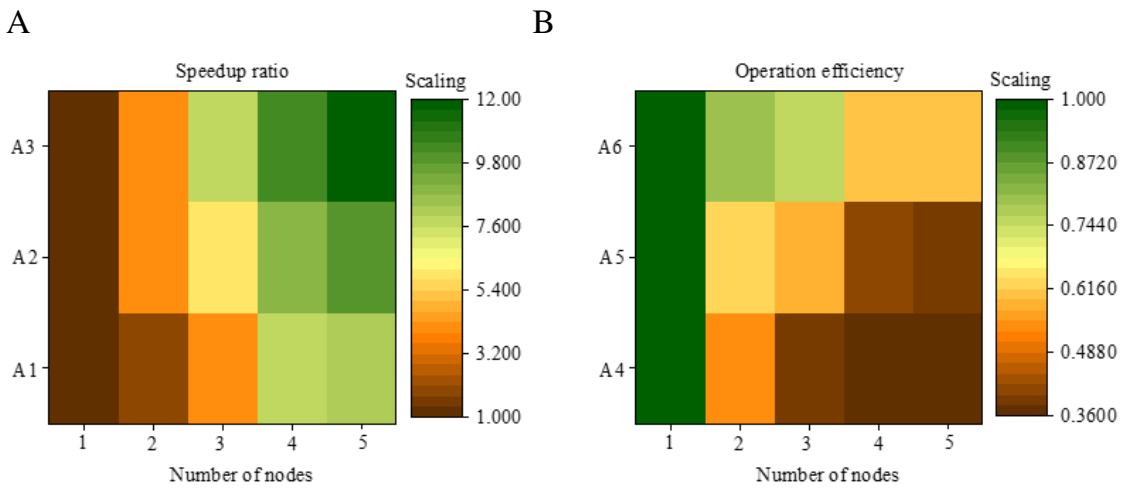

Figure. 9 Trajectory prediction results based on LSTM network.
(Note: A1-A3 are 10,000, 30,000, and 50,000 pieces of data, and A4-A6 are 20,000, 40,000, and 80,000 pieces of data, respectively.)

4. Conclusion

Investigating the hazard situation through field investigation or visual interpretation of radar images is with many shortcomings, such as threatening the personal safety, time-consuming and labor-intensive, and high professional quality requirements. To introduce the CNN technology, which is featured prominently in the intelligent interpretation of radar images, into the geologic hazard information extraction, a scientific and practical geologic hazard radar identification model is constructed, which performed better with its algorithm. The above research can provide a theoretical basis for the adoption of CNN technology in geologic hazard identification via radar image. Although the constructed geologic hazard radar identification model can better serve the adoption of CNN in geologic hazard identification, there are still some shortcomings. The constructed radar image sample library for geologic hazard contains a total of 1,316 radar scene images. Compared with other general radar image sample libraries, the number is small, which needs to be further expanded later. The constructed model data is based on Landsat 8, and the source is relatively single. Later, other data sources will be added, such as UAV images and China's high-resolution satellite images. It will continue to conduct in-depth research in these aspects.

Acknowledgements

1. This work was supported by Supported by SiChuan key research and development plan"geological hazard identification technologies of mountain tourist line based on big data"(No. 2020YFS0354)

2. This work was supported by Supported by Scientific Research Fund Project of Yunnan Provincial Department of Education "Earthquake relief asymmetric information game dynamics model in complicated landforms" (No. 2020J0381).

Compliance with Ethical Standards

Conflict of Interest: All Authors declare that they have no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent was obtained from all individual participants included in the study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

References

[1] Amerta I M S. The role of tourism stakeholders at Jasri tourism village development, Karangasem regency. International journal of social sciences and humanities, 2017, 1(2), pp. 20-28.

[2] Becken S, Khazai B. Resilience, tourism and disasters. Tourism and Resilience; Butler, RW, Ed; CAB International: Boston, MA, USA, 2017, pp. 96-102.
[3] Liu Y, Wu L. Geological disaster recognition on optical remote sensing images using deep learning. Procedia Computer Science, 2016, 91, pp. 566-575.
[4] Jiang W, Rao P, Cao R, Tang Z, Chen K. Comparative evaluation of geological disaster susceptibility using multi-regression methods and spatial accuracy validation. Journal of Geographical Sciences, 2017, 27(4), pp. 439-462.
[5] Deng S, Li W. Spatial case revision in case-based reasoning for risk assessment of geological disasters. Geomatics, Natural Hazards and Risk, 2020, 11(1), pp. 1052-1074.
[6] Fleming K, Abad J, Booth L, Schueller L, Baills A, Scolobig A, Petrovic B, Zuccaro G, Leone M F. The use of serious games in engaging stakeholders for disaster risk reduction, management and climate change adaption information elicitation. International Journal of Disaster Risk Reduction, 2020, 49, pp. 101669-101676.
[7] Bai E, Guo W, Tan Y. Negative externalities of high-intensity mining and disaster prevention technology in China. Bulletin of Engineering Geology and the Environment, 2019, 78(7), pp. 5219-5235.
[8] Wu C, Li X, Chen W, Li X. A review of geological applications of high-spatial-resolution remote sensing data. Journal of Circuits, Systems and Computers, 2020, 29(06), pp. 2030006-2030012.
[9] Chang P, Grinband J, Weinberg B, Bardis M, Khy M, Cadena G, Su M-Y, Cha S, Filippi C, Bota D. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. American Journal of Neuroradiology, 2018, 39(7), pp. 1201-1207.
[10] Liu B, Zhang Y, He D, Li Y. Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry, 2018, 10(1), pp. 11-18.
[11] Wang P, Zhang H, Patel V M. SAR image despeckling using a convolutional neural network. IEEE Signal Processing Letters, 2017, 24(12), pp. 1763-1767.
[12] Zhou Y, Wang H, Xu F, Jin YQ. Polarimetric SAR image classification using deep convolutional neural networks. IEEE Geoscience and Remote Sensing Letters, 2016, 13(12), pp. 1935-1939.
[13] Shen C, Lin H, Fan X, Chu Y, Yang Z, Wang J, Zhang S. Biomedical event trigger detection with convolutional highway neural network and extreme learning machine. Applied Soft Computing, 2019, 84, pp. 105661-105669.
[14] Can R, Kocaman S, Gokceoglu C. A convolutional neural network architecture for auto-detection of landslide photographs to assess citizen science and volunteered geographic information data quality. ISPRS International Journal of Geo-Information, 2019, 8(7), pp. 300-308.
[15] Xu Y, Qiu X, Yang X, Lu X, Chen G. Disaster risk management models for rural relocation communities of mountainous southwestern China under the stress of geological disasters. International Journal of Disaster Risk Reduction, 2020, pp. 101697-101676.
[16] Sameen M I, Pradhan B, Lee S. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. Catena, 2020, 186, pp. 104249-104253.
[17] Chen G P, Zhao J S, Yuan L, Ke Z J, Gu M, Wang T. Implementation of a Geological Disaster Monitoring and Early Warning System Based on Multi-source Spatial Data: A Case Study of Deqin County, Yunnan Province. Natural Hazards and Earth System Sciences Discussions, 2017, pp. 1-15.
[18] Li S, Wu J, Xu Z, Li L. Unascertained measure model of water and mud inrush risk evaluation in karst tunnels and its engineering application. KSCE Journal of Civil Engineering, 2017, 21(4), pp. 1170-1182.
[19] Xu X, Peng S, Yang F. Development of a ground penetrating radar system for large-depth disaster detection in coal mine. Journal of Applied Geophysics, 2018, 158, pp. 41-47.
[20] Guo W Y, Tan Y L, Zhao T B, Liu X M, Gu Q H, Hu S C. Compression creep characteristics and creep model establishment of gangue. Geotechnical and Geological Engineering, 2016, 34(4), pp. 1193-1198.
[21] El-Sawy A, Hazem E B, Loey M. CNN for handwritten arabic digits recognition based on LeNet-5. International conference on advanced intelligent systems and informatics: Springer, 2016, pp. 566-575.
[22] Lu S, Lu Z, Zhang Y D. Pathological brain detection based on AlexNet and transfer learning. Journal of computational science, 2019, 30, pp. 41-47.
[23] Ke H, Chen D, Li X, Tang Y, Shah T, Ranjan R. Towards brain big data classification: Epileptic EEG identification with a lightweight VGGNet on global MIC. IEEE Access, 2018, 6, pp. 14722-14733.
[24] Khan R U, Zhang X, Kumar R. Analysis of ResNet and GoogleNet models for malware detection. Journal of Computer Virology and Hacking Techniques, 2019, 15(1), pp. 29-37.

[25] Targ S, Almeida D, Lyman K. Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:160308029, 2016, pp. 1-7.
[26] Held B S. Positive psychology's a priori problem. Journal of Humanistic Psychology, 2018, 58(3), pp. 313-342.
[27] Singh S, Garg R, Mishra P. Performance optimization of MapReduce-based Apriori algorithm on Hadoop cluster. Computers & Electrical Engineering, 2018, 67, pp. 348-364.
[28] Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, Kamatani Y. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. Genome biology, 2019, 20(1), pp. 117-121.

Figures

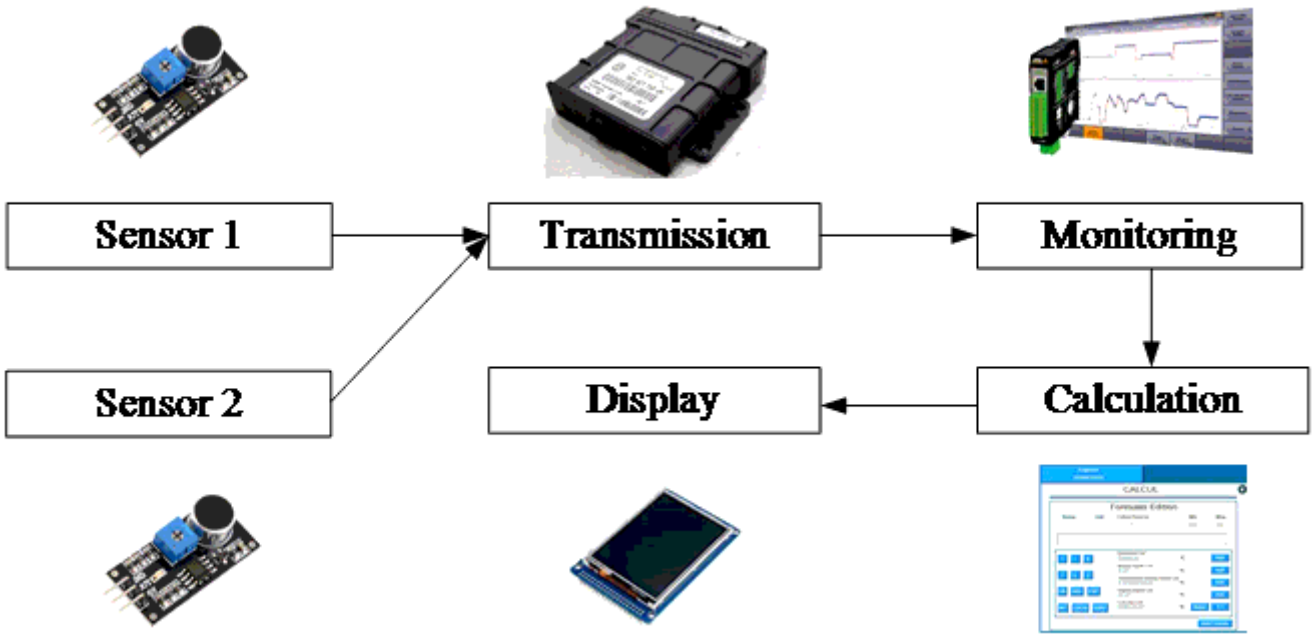

Figure 1

Traditional geologic hazard identification model.

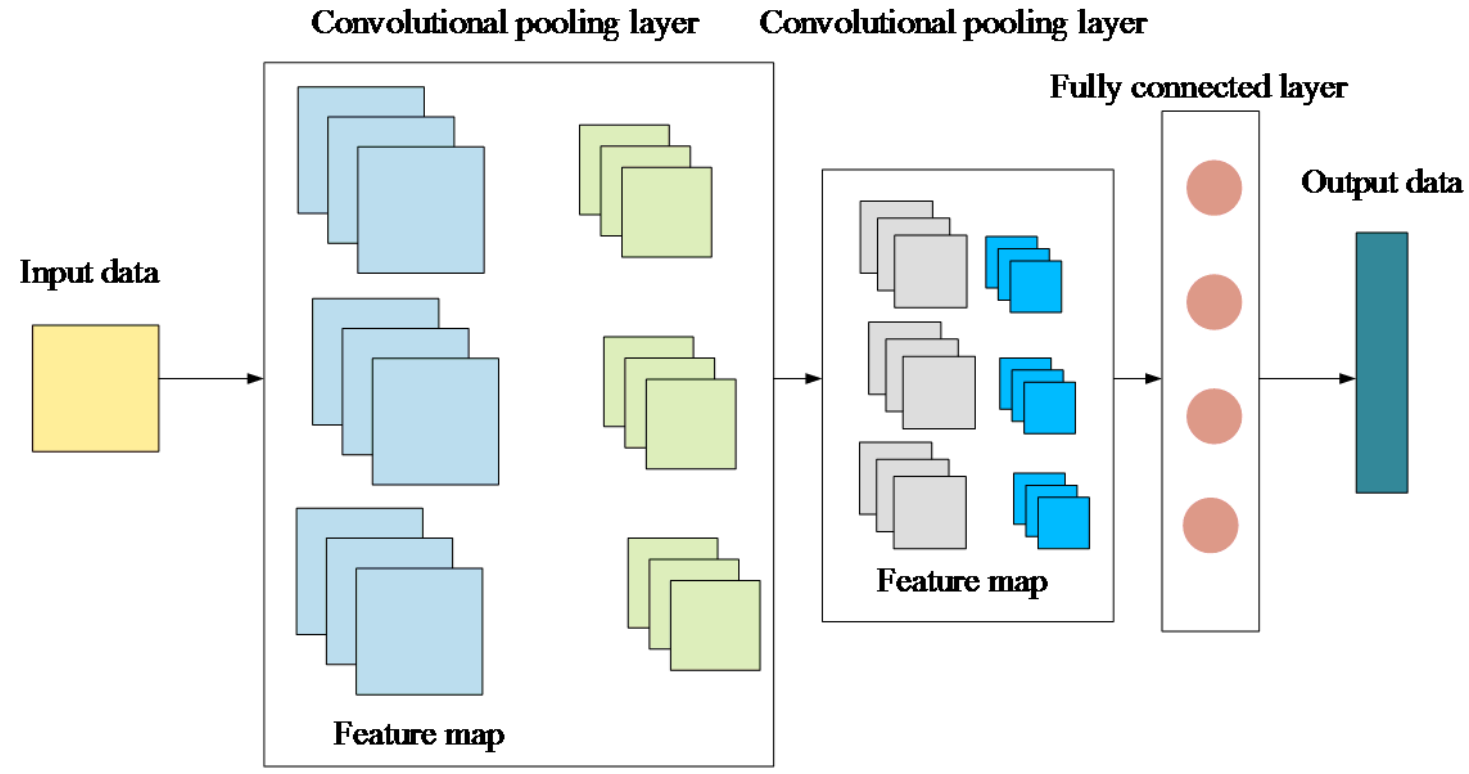

Figure 2

Schematic diagram of CNN structure.

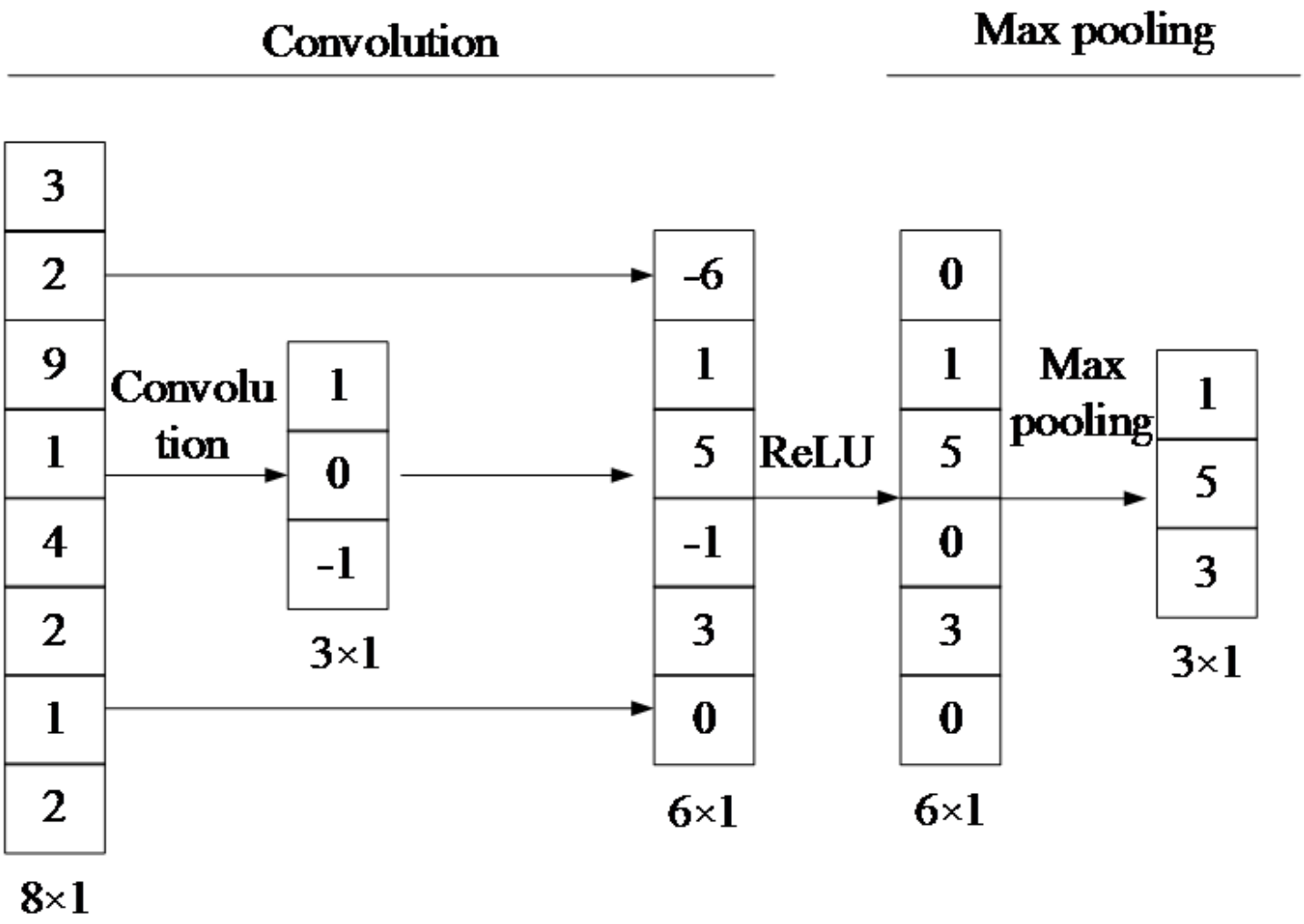

Figure 3

Schematic diagram of data pre-processing in CNN input layer.

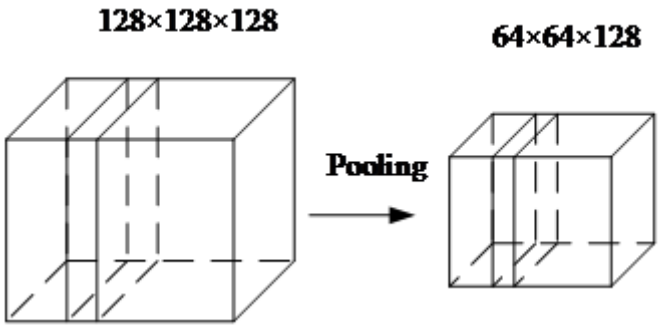

Figure 4

Schematic diagram of CNN pooling layer.

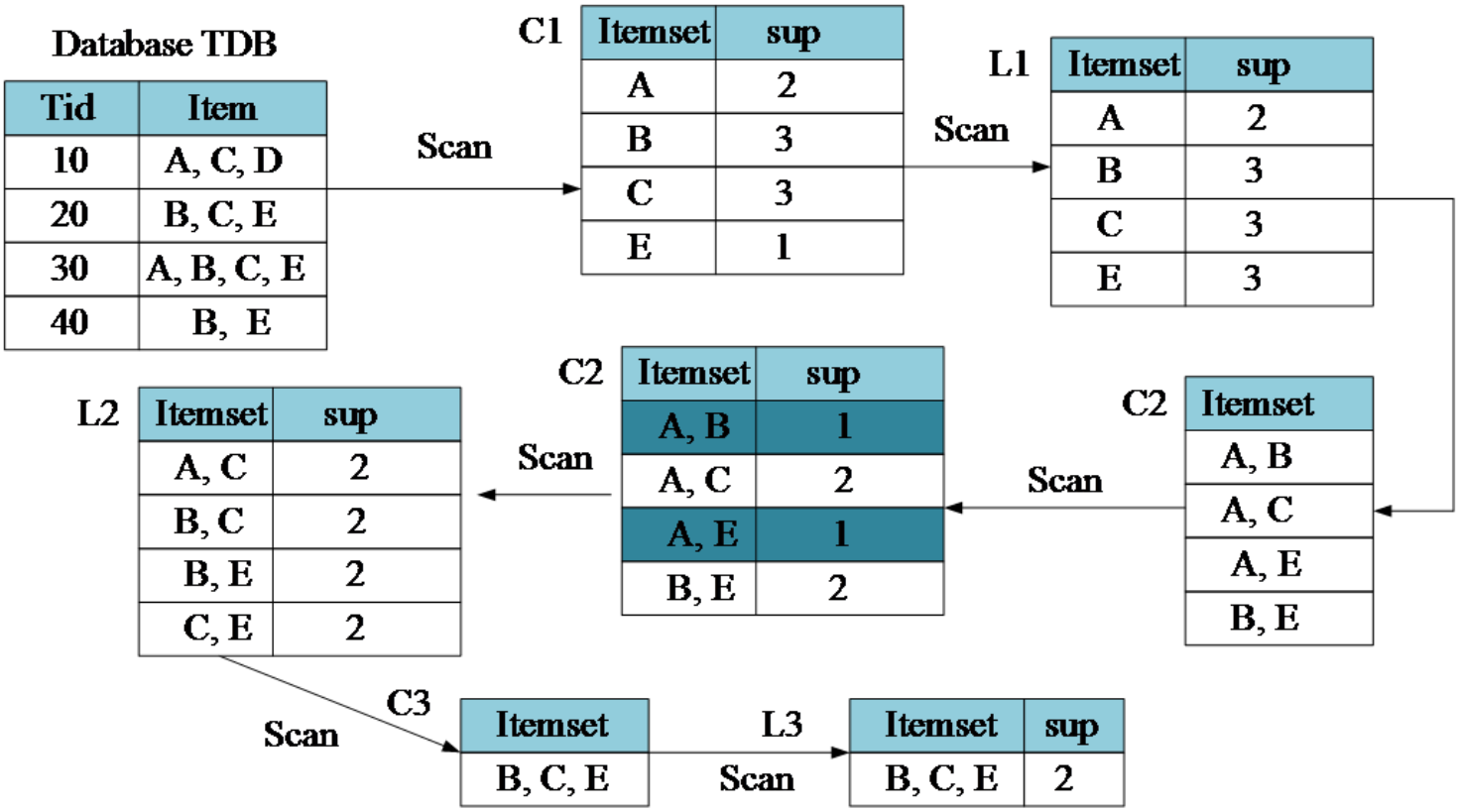

Figure 5

Flow of Apriori algorithm.

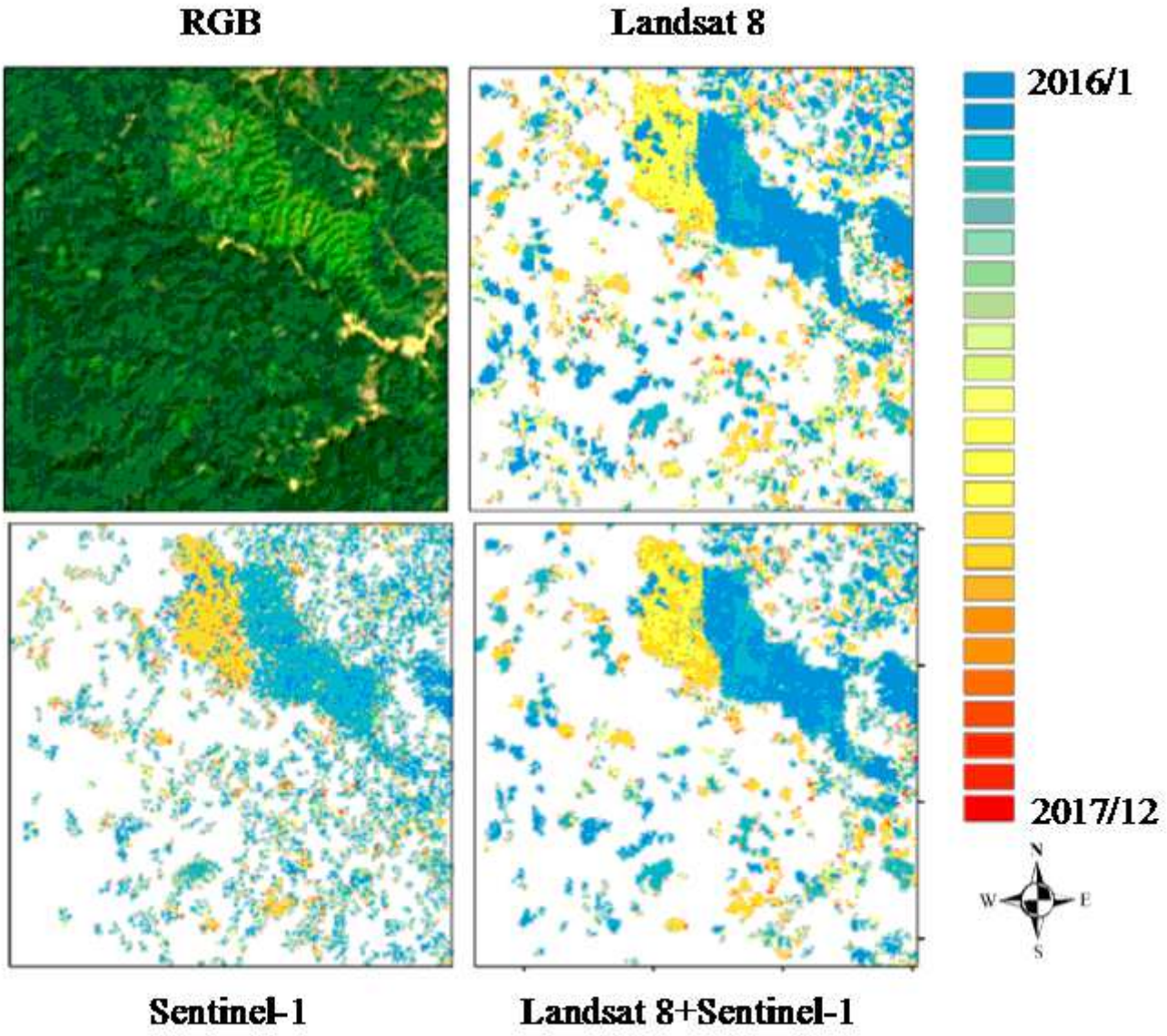

Figure 6

Sample radar data of Landsat 8 and Sentinel-1.

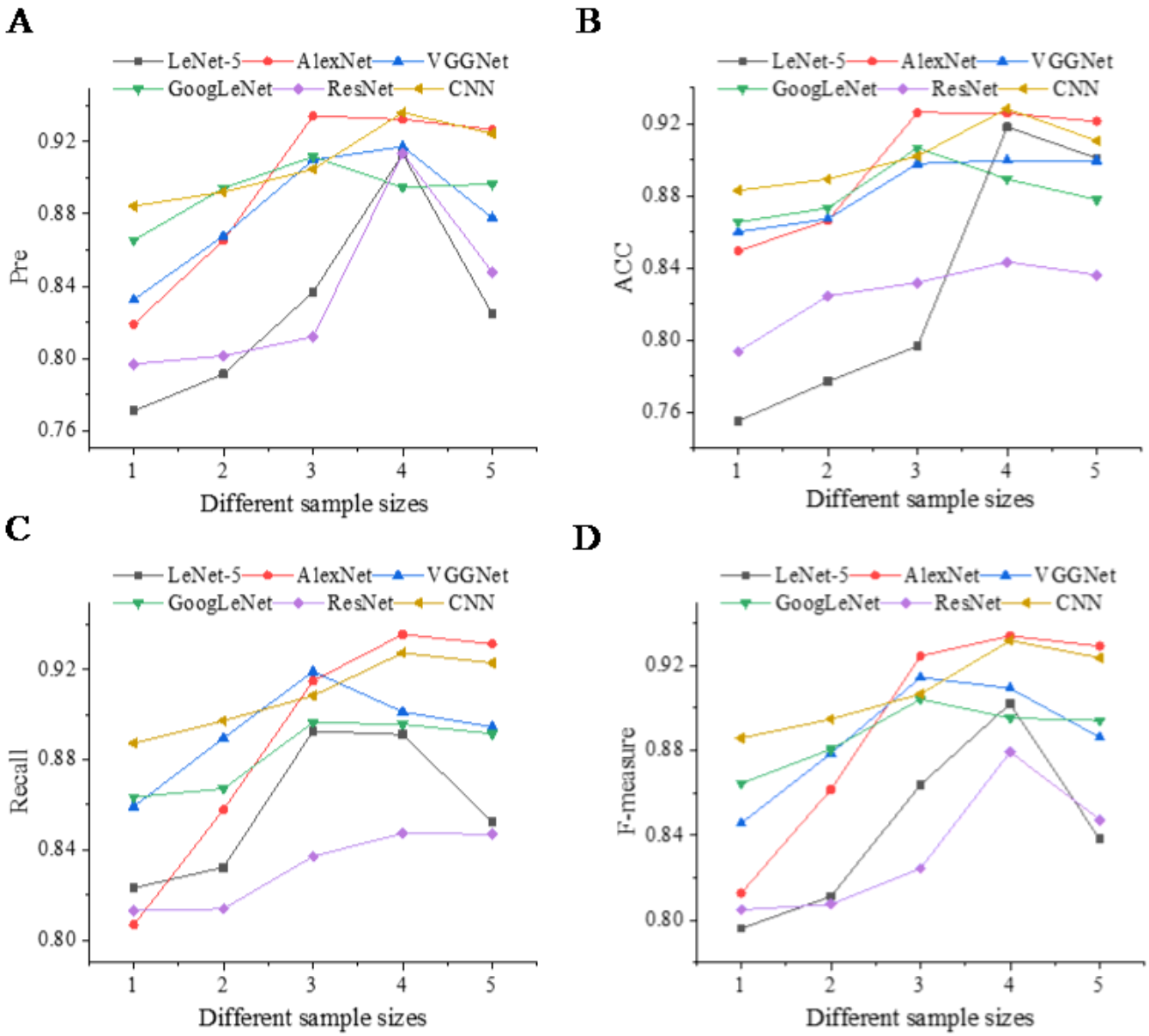

Figure 7

The effect of sample size on identification effect. (Note: 1-5 are five resolution sizes of 10×10, 50×50, 128×128, 256×256, and 512×512, respectively.)

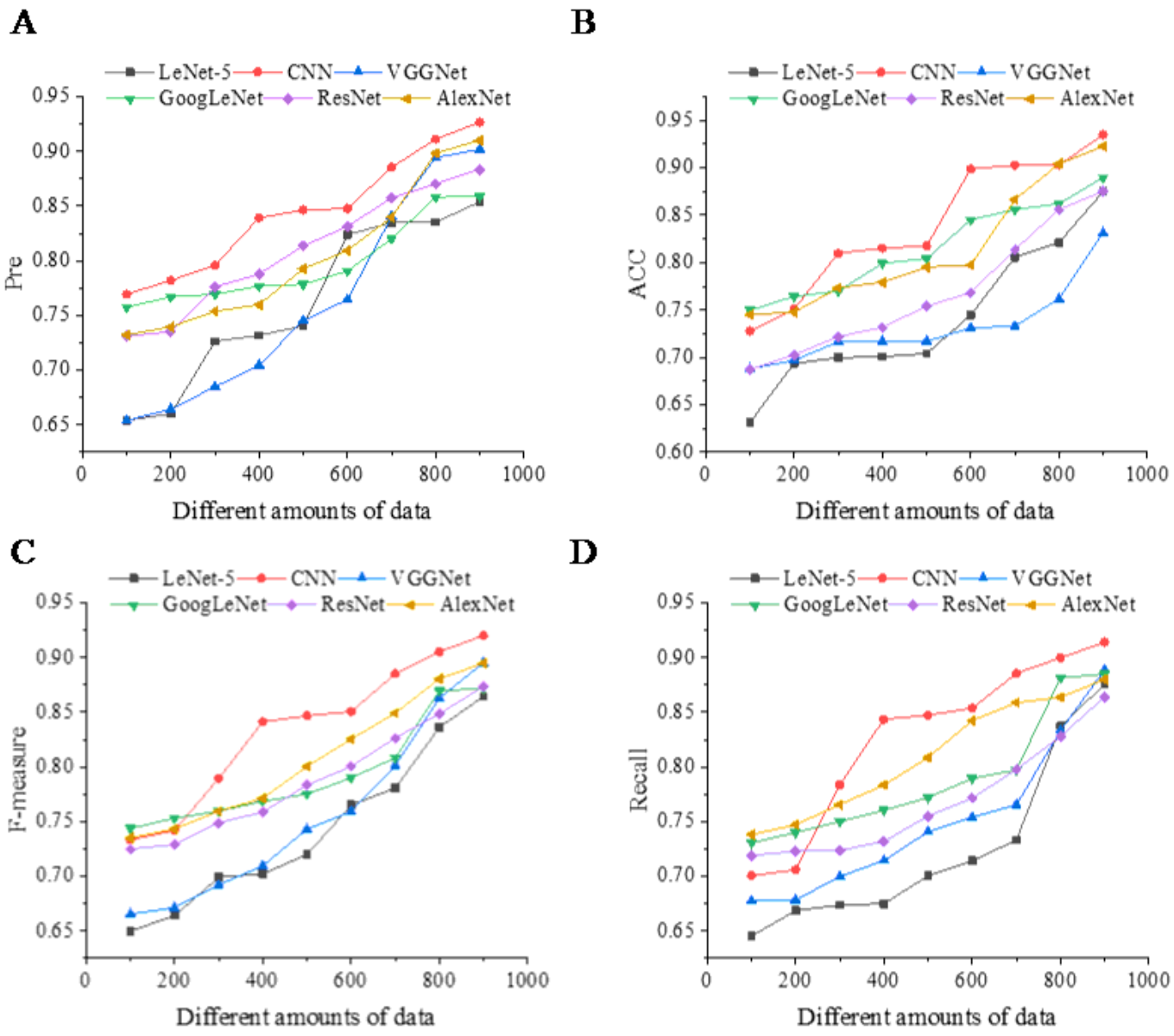

Figure 8

The effect of data size on identification effect.

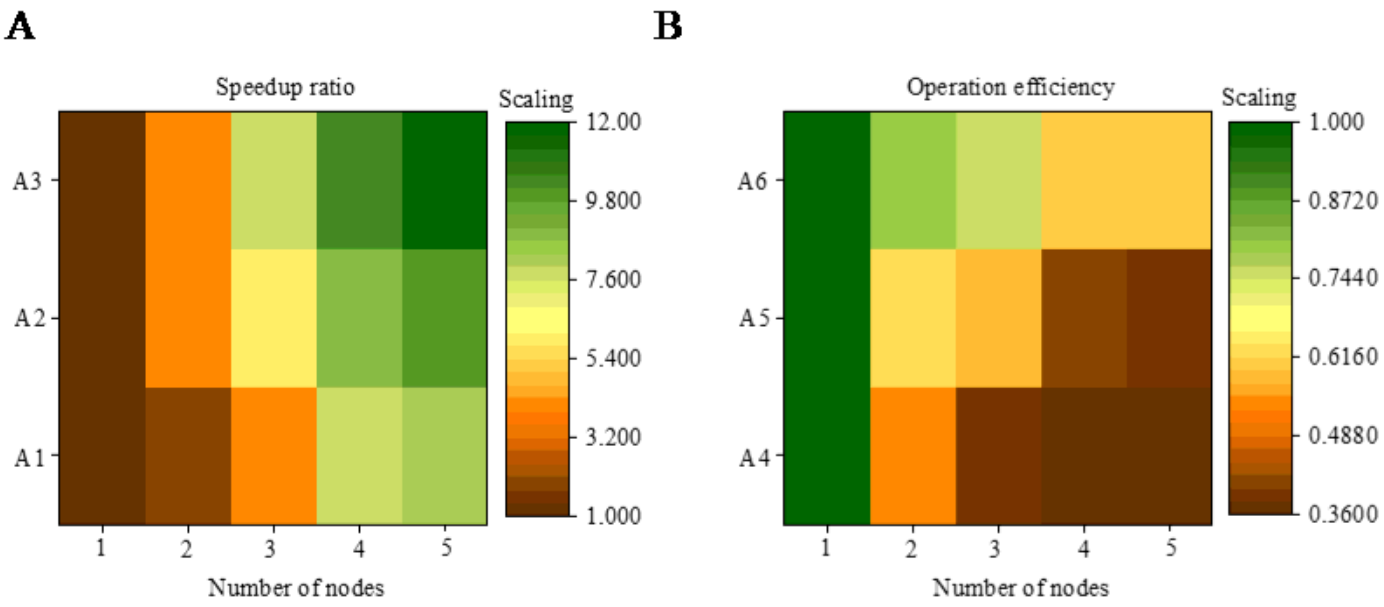

Figure 9

Trajectory prediction results based on LSTM network. (Note: A1-A3 are 10,000, 30,000, and 50,000 pieces of data, and A4-A6 are 20,000, 40,000, and 80,000 pieces of data, respectively.)