Urban built-up areas extraction by the multiscale stacked denoising autoencoder technique

Xiaofei Mi
Weijia Cao
Jian Yang
Zhenghuan Li
Yazhou Zhang
Qianjing Li
Zhensheng Sun
Yulin Zhan

Urban built-up areas extraction by the multiscale stacked denoising autoencoder technique

Xiaofei Mi,[a,b] Weijia Cao,[a] Jian Yang,[a,*] Zhenghuan Li,[c,d] Yazhou Zhang,[e] Qianjing Li,[f] Zhensheng Sun,[a,b] and Yulin Zhan[a]

[a]Chinese Academy of Sciences, Aerospace Information Research Institute, Beijing, China
[b]University of Chinese Academy of Sciences, Beijing, China
[c]Tsinghua University, School of Economics and Management, Beijing, China
[d]China Fortune Land Development Industrial Investment Co., Ltd., Beijing, China
[e]LuDong University, School of Resources and Environmental Engineering, Shandong, China
[f]Taiyuan University of Technology, College of Mining Engineering, Shanxi, China

Abstract. Stacked denoising autoencoder (SDAE) model has a strong feature learning ability and has shown great success in the classification of remote sensing images. However, built-up area (BUA) information is easily interfered with by broken rocks, bare land, and other features with similar spectral features. SDAEs are vulnerable to broken and similar features in the image. We propose a multiscale SDAE model to overcome these problems, which can extract BUA features in different scales and recognize the type of land object from multiple scales. The model effectively improves the recognition rate of BUA. The experimental results show that our algorithm can resist the disturbance information, and the classification accuracies are better than support vector machine, backpropagation, random forests, and SDAE. Then we investigate an application in Wuhan (China) metropolitan area analysis with the classification results of our algorithm. The range of the metropolitan area is 1.5-h isochronous circle calculated by Tencent map big data and is divided into three layers: core metropolitan area, subcore metropolitan area, and daily metropolitan. Finally, from the comprehensive statistical data and traffic data, we know that the Wuhan metropolitan area has a "target-shaped" distribution structure radiating outward from the core metropolitan area. It includes five metropolitan development corridors: Wuhan–Huanggang, Wuhan–Xiaogan–Suizhou, Wuhan–Ezhou–Huangshi, Wuhan–Xiantao–Tianmen, and Wuhan–Xianan–Chibi. The corridor is of great significance to the development of metropolitan areas.

1 Introduction

With the representative features of influence and radiation, the metropolitan area plays an increasingly important role in the field of city clusters.[1] The metropolitan area has the characteristic of urban agglomeration to some extent. With the development of rapid traffic and the expansion of commuter circle, the range of the metropolitan area is relatively larger.[2] Domestic and foreign research on the metropolitan area mainly focus on commuting distance and spatial distribution of built-up area (BUA).[3,4] Distance factors are mostly used to define the range of the metropolitan area, but there are no strict standards for travel time and spatial distance.[5–8] Zakaria[9] took Philadelphia as an example, studied the relationship between regional public transport or car accessibility and land use growth rate, and found that the growth rate of different areas in the urban area was inconsistent due to different accessibility. Liu et al.[10] consulted the transportation planning of highway and waterway (2002 to 2020) and skeleton highway network planning (2002 to 2020) of Hubei Province proposed the time–distance accessibility model, and established

---

1-h high-accessibility circle (0 to 75 km), 2-h medium accessibility circle (75 to 150 km), and 3-h low-accessibility circle. With the development of big data analysis and collection methods, mobile phone signaling data and traffic data have been widely applied in commuting analysis. It realizes the cognition of residents' travel needs and travel characteristics from the city microscale.

BUA is a relatively concentrated area that contains buildings, public facilities, and urban roads within the urban administrative area.[11] We define BUA for which the surface is predominantly impervious, including all nonvegetative, nonwater, nonsoil, human-constructed elements (e.g., roads and buildings). BUA is an objective reflection of urban construction and development in regional distribution and indicates the scale and size of construction land in different periods of urban development.[12] Spatial information of BUA is fundamental to a better understanding of the development direction of the metropolitan area, the inter-relationship between cities, and the radiation effect of core cities on surrounding cities.[13] It is of great significance for guiding and evaluating the development of the metropolitan area.[14]

Many BUA extraction methods using remote sensing images have been proposed.[15,16] They can be mainly divided into three categories, including pixel-based, object-based, and deep learning-based methods. The pixel-based method is a popular method to classify remote sensing images, given its simplicity and high efficiency. However, the classification results display a "salt and pepper" effect. To overcome this problem, the object-based method has become a mainstream method in land-use/land-cover application recently. However, the segmentation scale is a key problem to the object-based method in the face of different data and application scenarios. Deep learning method is proposed to improve these performances with its perfect fitting ability. Its neural network has strong expression property to imitate various complex models, and thus it can be widely applied to the land use/land cover.

Stacked denoising autoencoder (SDAE) is a typical deep learning method and works in much the same way as stacking restricted Boltzmann machine in deep belief networks or ordinary autoencoders.[17,18] It learns to recover the corrupted data with the help of an unsupervised pretraining procedure that initializes the neural network. Then it seeks to be trained over the entire neural network using supervised learning to recognize the moving target.[19,20,21] Zhang et al.[22] extracted the spectral, spatial, and texture features for each object and put all features into stacked autoencoder or SDAE network, and then got the parameters of the network. The classification result is better than that of "linear" support vector machine (SVM) model and radial basis function (RBF) SVM model. Han et al.[23] used SDAE to predict human eye fixations in two steps. He used center patch and its surrounding patches to represent the features, developed model to learn feature from raw image data under an unsupervised manner, and then captured the intrinsic mutual patterns as the feature contrast and integrated them for final saliency prediction. Li[24] used SDAE and Softmax model to solve the problems of automatic feature extraction and dimension reduction in Braille recognition. The SDAE performs better than the traditional feature extraction algorithms and Softmax has a better performance than multilayer perceptron and RBF when they perform with SDAE. Zhang[25] constructed the detection vector of the center pixel based on the center pixel and its neighboring pixels and used the SDAE model to classify the land cover based on GF-1 image. The classification result is better than that of traditional SVM and backpropagation (BP) network. SDAE has been widely applied to feature learning in many other fields, such as denoising and target recognition.[26–31,32] The spatial distribution of BUA concentrates on distribution, consistent types, and structures. It is easy to be disturbed in the process of classification by broken rocks, bare land, and other features with similar spectral features. SDAE model has a strong feature learning ability, but it lacks spatial scale features. When encountering the phenomenon of "different object with the same spectra feature and same object with the different spectra feature in remote sensing image," the classification ability is limited.

In this paper, we develop a new BUA extraction method based on SDAE, and the BUA results extracted by this method are applied to the analysis of metropolitan area. First, the recognition of the same land object will have different results in different scales. It manifests that the spatial pattern of land objects is significantly different at different scales.[33] We will analyze the spatial scale features of BUA, generate multiscale hierarchical structure features, and integrate the learning ability of SDAE model. Then we propose a multiscale stacked denoising autoencoder (MSDAE) model to learn the features and extract BUA from multiple scales. It improves

the classification ability of BUA. Second, taking Wuhan city for example, we divide the commuting isochronous circles into 0.5-h isochronous circle (0.5 h), 1-h isochronous circle (1 h), 1.5-h isochronous circle (1.5 h) base on Tencent map big data. We comprehensively analyze the clustering degree of BUA, population density, urban traffic, and corridor in this area.

The organization of paper is as follows: Sec. 2 gives an introduction of region and data; Sec. 3 introduces the SDAE and describes the proposed method in detail; Sec. 4 presents the extraction result; Sec. 5 presents precision evaluation; following that, based on the result of BUA extracted using MSDAE method, metropolitan area analysis are given in Sec. 6. Finally, the conclusion is drawn in Sec. 7.

2 Study Area and Data

2.1 Study Area

Wuhan is located in the east of Jianghan plain and in the middle of the Yangtze river, 113°41′E–115° 05′ E, 29°58 ′N–31°22′N. The terrain in north is higher than in south, and it is mostly flat in middle. The average elevation of this area is 23.3 m with the same wave rolling hills and plains geomorphology. Wuhan is the capital of Hubei Province, the only subprovincial city and megacity in the six central provinces, the central city of central China, and an important industrial base, scientific and educational base, and comprehensive transportation hub in China.

Taking a central city as the center, the regional accessibility can well explain the radiation capacity and the degree of connection of the central city to the surrounding areas in different directions.[34]

Based on population heat map and real traffic flow from Tencent map big data, we get the center of the city, calculate the distance from the center to the furthest point; it, respectively,
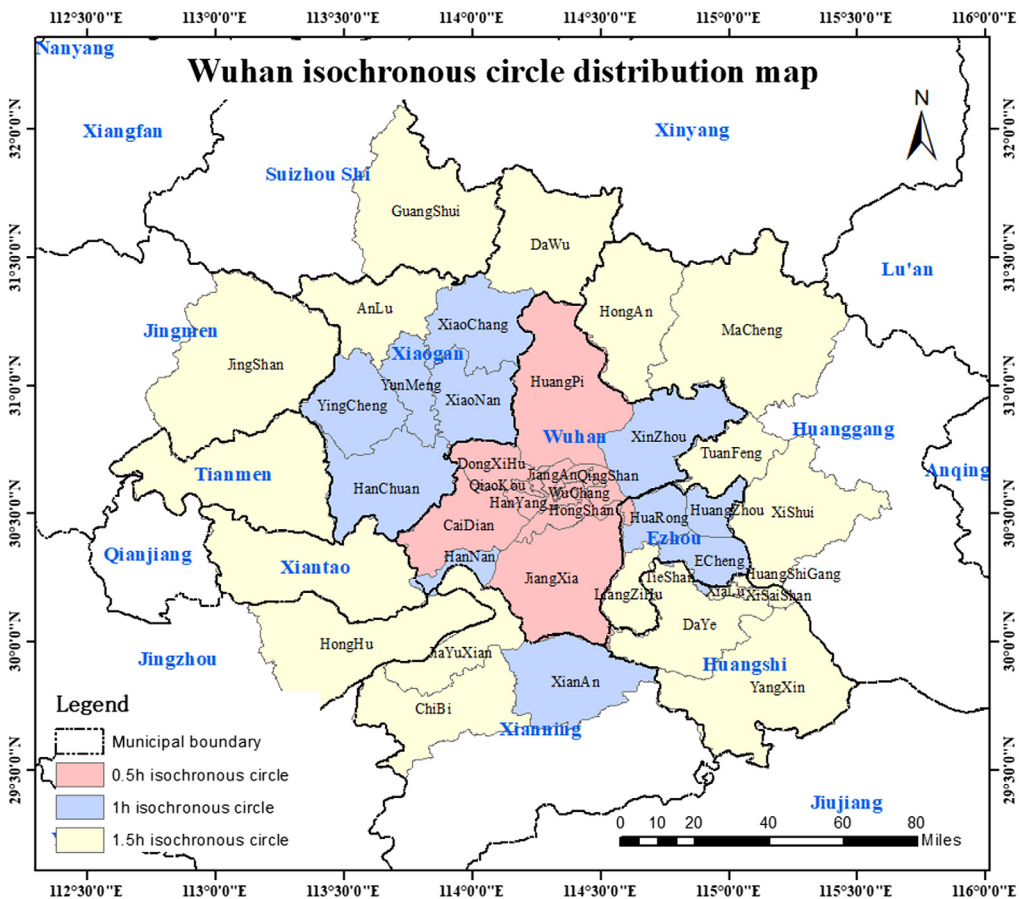

**Fig. 1** Wuhan isochronous circle distribution map.

Table 1 Area covered by isochronous circle statistics.

Isochronous circle	Districts and counties
0.5 h	Huang Pi, Dong Xihu, Hong Shan, Jiang Xia, Cai Dian, Qing Shan, Han Yang, Wu Chang, Jiangan, Jianghan, Qiao Kou
1.0 h	Xin Zhou, Xiao Nan, Xiao Chang, Yun Meng, Ying Cheng, Han Chuan, Huang Zhou, Hua Rong, E Cheng, Xian An, Han Nan, Tie Shan
1.5 h	Jing Shan, Tuan Feng, An Lu, Tian Men, Xi Shui, Da Ye, Xi Sai Shan, Liang Zi Hu, Jia Yu Xian, Xian Tao, Xia Lu, Huang Shi Gang, Guang Shui, Da Wu, Hong An, Ma Cheng, Chi Bi, Yang Xin, Hong Hu

takes 0.5, 1, and 1.5 h (considering the complexity of road condition during the day, we use the night traffic flow condition of travel time). By overlaying the administrative zoning map with the furthest distance, we get the isochronous circles based on administrative zoning map (see Fig. 1). In this paper, we define 1.5 h as the boundary of the metropolitan area. The metropolitan area covers 42 districts and counties, involving 10 cities. Among them, 0.5 h covers 11 districts and counties, 1 h covers 12 districts and counties, and 1.5 h covers 19 districts and counties, as shown in Table 1.

2.2 Data and Preprocess

The metropolitan area of Wuhan is taken as the research area; all available GF-1 WFV images data with cloud cover less than 10% are chosen for inclusion in this study and acquired in April 2018 with a spatial resolution of 16 m are selected. The preprocess includes geometric correction and mosaic and projection transformation. The result of processing is shown in Fig. 2.
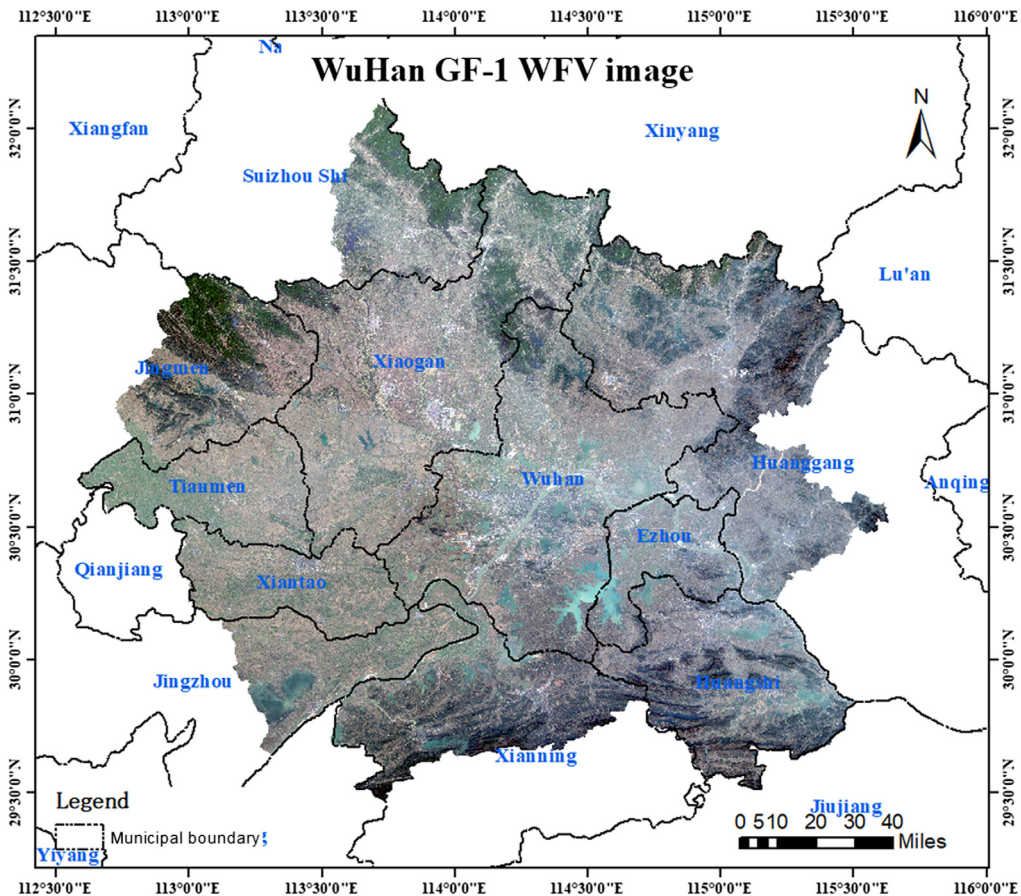
Fig. 2 GF-1 image of the study area.

3 Multiscale Stacked Denoising Autoencoder Method

In this section, we elaborate on the proposed MSDAE to extract BUA. First, we introduce the related work, namely, SDAE algorithm and its characteristics. Next, the proposed method is described in detail.

3.1 *Stacked Denoising Autoencoder*

The architecture of SDAE is divided into two steps. The first is feature learning, which is a process of unsupervised learning. The second step is optimization of network parameter, which is a process of supervised learning (Fig. 3).

  The basic building block of an SDAE is denoising autoencoder (DAE), which is one variant of the standard autoencoder.[35,36] Autoencoder can learn to recover the data from the corresponding corrupted input data. SDAE allows us to build a deep network to use denoise feature as an unsupervised objective to guide the learning of useful higher level representations.[37] As mentioned in Vincent's research, the autoencoder framework comprises two parts: encoder and decoder. The DAE is trained to reconstruct a clean "repaired" input from a corrupted version of it. Before encoding, the initial input x into $\tilde{x}$ is done by means of a stochastic mapping $\tilde{x} \sim qD(\tilde{x}|x)$ in which some elements of x is forced to be zero randomly (masking noise). Then the encoder procedure is provided a nonlinear affine mapping function $f_\theta(x)$, which transforms the corrupted vector into a hidden representation by the following equation:

$$y_i = f_\theta(\tilde{x}_i) = \text{sigm}[W^{(l)}\tilde{x}_i + b^{(l)}].\qquad(1)$$

  Its parameter $\theta = \{W, b\}$, where W is the weight matrix and b is an offset vector. The activation function sigm is set to sigmoid function, where $\text{sigm} = 1/(1 + e^{-x})$. A decoder is the process where the hidden representation y_i is mapped back to a reconstructed vector z_i in a similar equation:

$$z_i = g_\theta(y_i) = \text{sigm}[W^{(l)}y_i + b^{(l)}].\qquad(2)$$

  To meet criteria of feature representation, features in the data can be learned by minimizing the reconstruction error of the loss function. To emphasize on the corrupted dimensions, the weights are set differently among all components of the input. The corrupted dimensions is emphasized, and the squared loss yields:

$$L_{2,\alpha}(x,z) = \alpha\left[\sum_{j\in\kappa(\tilde{x})}(x_j - z_j)^2\right] + \beta\left[\sum_{j\notin\kappa(\tilde{x})}(x_j - z_j)^2\right],\qquad(3)$$

where $\kappa(\tilde{x})$ denotes the indices of the components of x that were corrupted. The weight α denotes the reconstruction error on components that were corrupted, and β denotes those that were left untouched; α and β are considered hyperparameters.
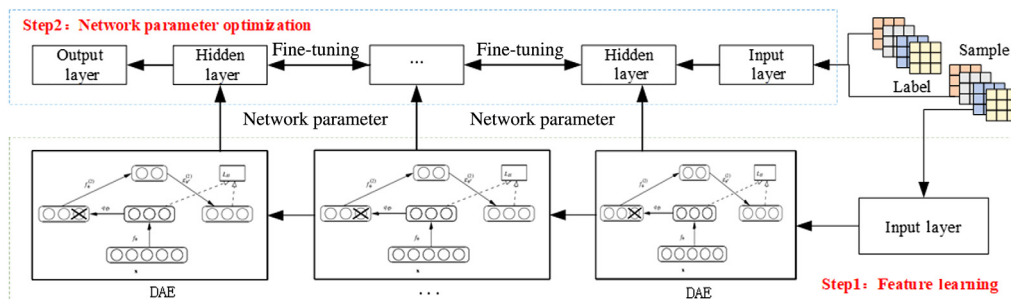

Fig. 3 The architecture of SDAE.

Finally, a feedforward neural network (FFNN) classifier can be added to the end of the deep neural network to form a complete SDAE model for image classification. Hidden layer network structure of FFNN is the same as hidden layer structure constructed of the SDAE. During training, the network parameters obtained by SDAE training are taken as the optimal initialization parameters of FFNN, and labeled samples are used to train the model. FFNN network uses error propagation mechanism, according to the difference between the output and the label, BP algorithm is used to fine-tune network parameters until convergence. The parameters of all layers are well tuned using the stochastic gradient descent algorithm.[23,38,39]

3.2 Multiscale Stacked Denoising Autoencoder

Many research results have shown that the scale is a critical factor in remote sensing image classification. Land objects in remote sensing images are complex and broken, so the land object needs to be recognized from different scales. To tackle this problem, we propose a new method to recognize the type of land object from multiple scales, which is called MSDAE. The mode contains three parts: multiscale training, multiscale classification, and multiscale results merging. The architecture of the model is shown in Fig. 4.

To accomplish this, we first collect two types of samples: built-up sample and nonbuilt-up sample. Then, we crop each sample with 3×3 pixels corresponding to scale 1, 7×7 pixels corresponding to scale 2, 15×15 pixels corresponding to scale 3, 25×25 pixels corresponding to scale 4. The difference between the four scales is that scale 1 is used to determine the land object type of the center pixel by the vector composed of the center pixel and its eight surrounding points. Scales 2, 3, and 4 are used to determine the type of land object by patch. Therefore, from this perspective, scale 1 is pixel-based classification and scales 2 to 4 are patch-based classification. Therefore, the dimensionality of input vectors is different at different scales, so the configuration employed for MSDAE is shown in Table 2. In the training stage, the main aim is to train the MSDAE model by the layer-wise pretraining and supervised fine-tuning.

For each input image to be classified, we test it in the four scales separately. In every scale, there is no overlap in each direction among patches. Then we can obtain the result of
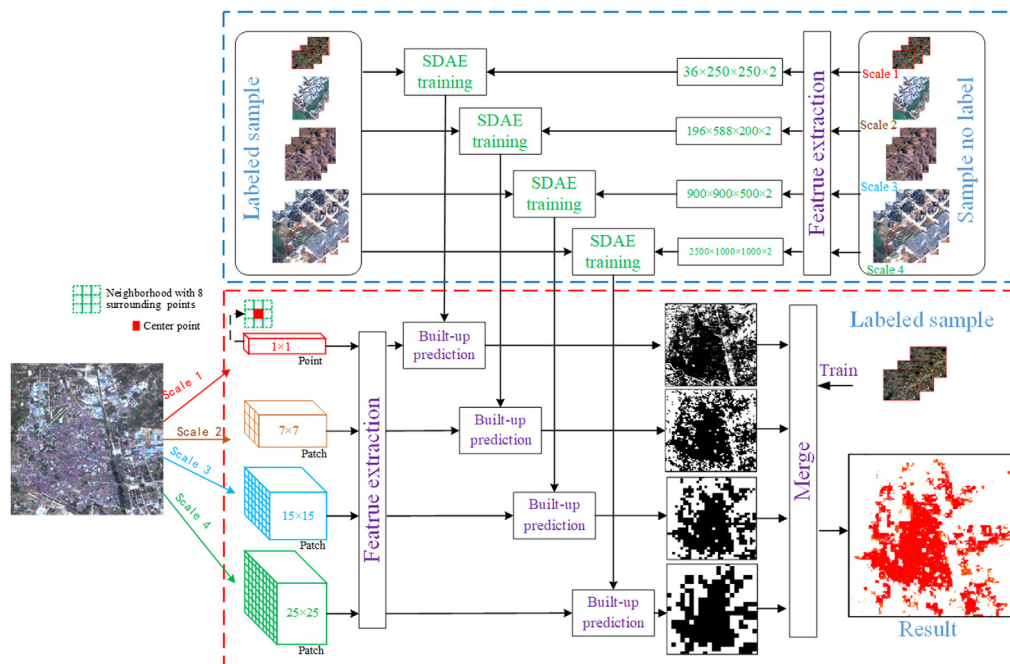

Fig. 4 The architecture of MSDAE.

Table 2 List of hyperparameter for MSDAE.

Hyperparameter	Description	Scale 1	Scale 2	Scale 3	Scale 4
nHLay	Number of hidden layers	{1,2}	{1,2}	{1,2}	{1,2}
nHUnit	Number of units per hidden layer	{250,250}	{588,200}	{900,500}	{1000,1000}
lRate	Fixed learning rate for unsupervised pretraining	0.1	0.1	0.01	0.01
lRateSup	Fixed learning rate for supervised pretraining	0.1	0.1	0.01	0.01
nEpoq	Number of pretraining epochs	{10,10,50}	{10,10,50}	{10,10,100}	{10,10,100}
v	Corrupting noise level	0.01	0.1	0.3	0.3

classification in each scale. The final result is calculated by the results at different scales, as shown in Eq. (4), where $L_{i,j}$ is the label of land object type of pixel point (i, j), $L_{i,j}^{S_n}$ is the label of object type of pixel point (i, j) in scale $n (n = 1,2,3,4)$, $L_{i,j}^{S_n} = 1$ is the BUA, $L_{i,j}^{S_n} = 0$ is the non-BUA; a, b, c, d represent the weight coefficients of the four scales respectively. The labeled sample points are used for logistic regression analysis, where L is the logistic function and p is the constant parameter. Finally, the final BUA extraction result is obtained from the classification results of the four scales, which is simple but effective.

$$L(aL_{i,j}^{S_1} + bL_{i,j}^{S_2} + cL_{i,j}^{S_3} + dL_{i,j}^{S_4} + p) = L_{i,j}. \tag{4}$$

4 Built-Up Area Extraction Result

The task in our experiments was to classify all pixels in images into two categories: built-up and nonbuilt-up using our model MSDAE. Experiments were conducted on GF-1 WFV image data that cover the metropolitan area of Wuhan [Fig. 5(a)], the size of the data is 20373×16376 and the data contain four spectral bands, which represents blue, green, red, and near-infrared in order. Using ROI tool from ENVI, 14,557 sample points of BUA and 72,206 sample points of non-BUA are selected on images. About 60% of them were randomly selected as training samples and 40% as test samples. The samples of MSDAE are made by expanding to 3×3, 7×7, 15×15, 25×25 slices with the pixel sample as the center. The model is trained and then is used to detect the image. The classification results are shown in Fig. 5. The result of SDAE-pixel-based [Fig. 5(b)] obviously has more noise than the result of MSDAE proposed in this paper. The noise is caused by unused land, ridges, rocks, and other similar land object.
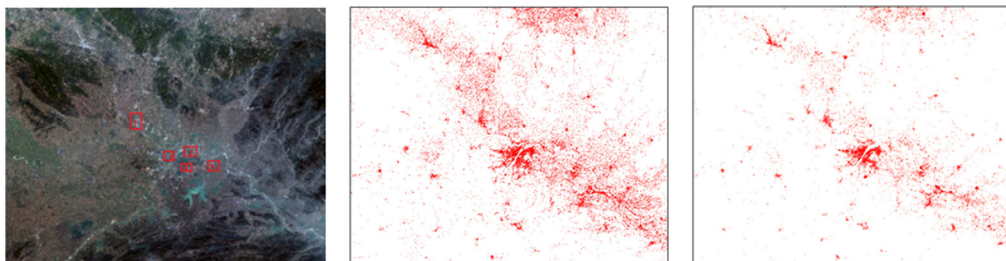
Fig. 5 (a) Image of GF-1, results of (b) SDAE-pixel-based and (c) MSDAE.

As shown in the Fig. 5, the detection of BUA from multiple scales can reduce the interference of other land objects.

5 Precision Evaluation

5.1 *Comparisons with Single-Scale Result*

In this paper, five regions [Fig. 5(a): 1–5] are selected to compare the classification results of scale 1, scale 2, scale 3, scale 4, and final results. As shown in Fig. 6, the classification result of scale 1 is of good detail, but the pepper-and-salt noise is obvious. The classification result of scale 2 is "granulated," but pepper-and-salt noise still exists. In the classification result at scale 3, the main part of the target object is highlighted and the noise is reduced. At scale 4, target object are more aggregated, but false alarm is amplified. As can be seen from the figure, the clustering degree of final result is higher and the noise is less. Through multiscale to confirm the type of land object type, we find that MSDAE can reduce the interference of similar features from other land object. R1 shows that both sides of the river are misclassified as BUA in scales 1 and 2 but not in scales 3 and 4. The misclassified land object is soil land, so through the fusion of the results from the four scales, misclassified information is excluded in the final result. R2 shows that the boundary of the BUA is gradually clear and the noise is reduced from scale 1 to scale 4. We get the BUA result with clear boundary and less noise by MSDAE. R3 shows that buildings are densely distributed and adjacent to other land objects, forming an obvious dividing line in the result of scale 1. It is made of bare land, but the BUA information of the result based on MSDAE is prominent. R4 and R5 show that BUA distributes sparsely, and the feature of BUA is similar to the feature of surrounding land object. The disturbance information of bare land, rock, and soil is identified by multiple scale discrimination, and the boundary of BUA in the final result is clear.

Precision, accuracy, recall, missing alarm, and false alarm are selected as evaluation indicators to evaluate and compare the classification results of scales 1 to 4. The precision, accuracy, recall, missing alarm, and false alarm are calculated by Eq. (5):

Fig. 6 Comparison diagram of results from different scales and MSDAE.

Table 3 Comparison table of classification accuracy from four scales and MSDAE.

	Scale 1	Scale 2	Scale 3	Scale 4	MSDAE
Precision	0.809	0.807	0.773	0.714	0.873
Accuracy	0.844	0.859	0.825	0.782	0.892
Recall	0.887	0.934	0.905	0.918	0.910
Missing Alarm	0.113	0.066	0.096	0.082	0.090
False Alarm	0.191	0.193	0.2275	0.286	0.127

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn},$$

$$\text{Precision} = \frac{tf}{tp + fp},$$

$$\text{Recall} = \frac{tp}{tp + fn},$$

$$\text{Missing alarm} = \frac{fn}{tp + fn},$$

$$\text{False alarm} = \frac{fp}{tn + fp}, \tag{5}$$

where tp is true positive, fp is false positive, fn is false negative, and tn is true negative. These values can be calculated by confusion matrices. As shown in Table 3, precision of MSDAE result is 0.873, which is best. Accuracy of scale 2 is 0.859, which is similar to accuracy of MSDAE. Recall of scales 3 and 4 is higher, which is related to its detection window. The larger the detection window, the higher is the probability that it will hit the target. Conversely, false alarm is higher. To sum up, MSDAE is superior to the results of single scale.

5.2 Comparisons with Other Methods

To further investigate and evaluate the performance of our framework, we compare the results of the classifications based on SVM, BP, random forest (RF), and SDAE. These methods are robust and widely used for land cover classification.[15,14] In our evaluation, overall accuracy (OA), user's accuracy (UA), producer's accuracy (PA), F1 score, intersection over union (IOU) are used to assess the quantitative performance from SVM, BP, RF, SDAE, and MSDAE. The F1 score is calculated by Eq. (6). IOU is the value of the intersection of prediction and ground-truth regions over their union, as shown in Eq. (7).

Note that all models were implemented on the same training dataset and test dataset. Table 3 compares the classification accuracies for five models from five evaluation indices. MSDAE consistently provided better results than other models and reached high scores (89.20% OA, 87.32% UA, 90.96% PA, 89.10% F1 score, and 80.35% IOU), which indicates that the MSDAE performed well on BUA extraction. The MSDAE clearly outperforms the SDAE by about 5% in the OA, about 7% in the UA, about 2% in the PA, about 5% in the F1 score, and about 7% in the IOU, respectively. This shows that the multiscale model is superior to the single-scale model. The proposed MSDAE can achieve a better performance (Table 4).

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \tag{6}$$

$$IOU = \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall} - \text{precision} \times \text{recall}}. \tag{7}$$

Table 4 Comparison table of classification accuracies from different models.

	OA (%)	UA (%)	PA (%)	F1 score (%)	IOU (%)
SVM	81.81	81.77	80.44	81.10	68.21
BP	80.49	80.97	78.18	79.55	66.04
RF	82.61	81.86	82.42	82.14	69.69
SDAE	84.38	80.93	88.70	84.64	73.37
MSDAE	89.20	87.32	90.96	89.10	80.35

6 Metropolitan Area Analysis

To study the aggregation and spatial distribution of BUA in the metropolitan area of Wuhan, the extraction result of BUA is processed; the postprocess includes the small patch removing and hole filling and cutting. The processed result is shown in the Fig. 7. The figure shows that the BUA concentrates in the center of the city and disperses outward. The metropolitan Wuhan was divided as core metropolitan area, subcore metropolitan area, and daily metropolitan area following the 0.5, 1, and 1.5 h. Taking the circle as research object, we compare and analyze the area of BUA and population of the administrative region between different circles.

According to the statistics, the 27.2714 million of permanent residents gathered in 49,890.29 km^2 metropolitan Wuhan in 2016. There were 56,439 km^2 BUA in the metropolitan area of Wuhan reported by our result. Proportion of district area, BUA, permanent resident population, and the population density are shown in Table 5. It indicates that the core metropolitan area took the 13.63% of whole metropolitan, in which the BUA took 44.02% and 32.35% of permanent residents gathered in with the population density of 1295 people/km^2; proportions in the subcore metropolitan area showed relatively balanced with about 20.41% proportion of
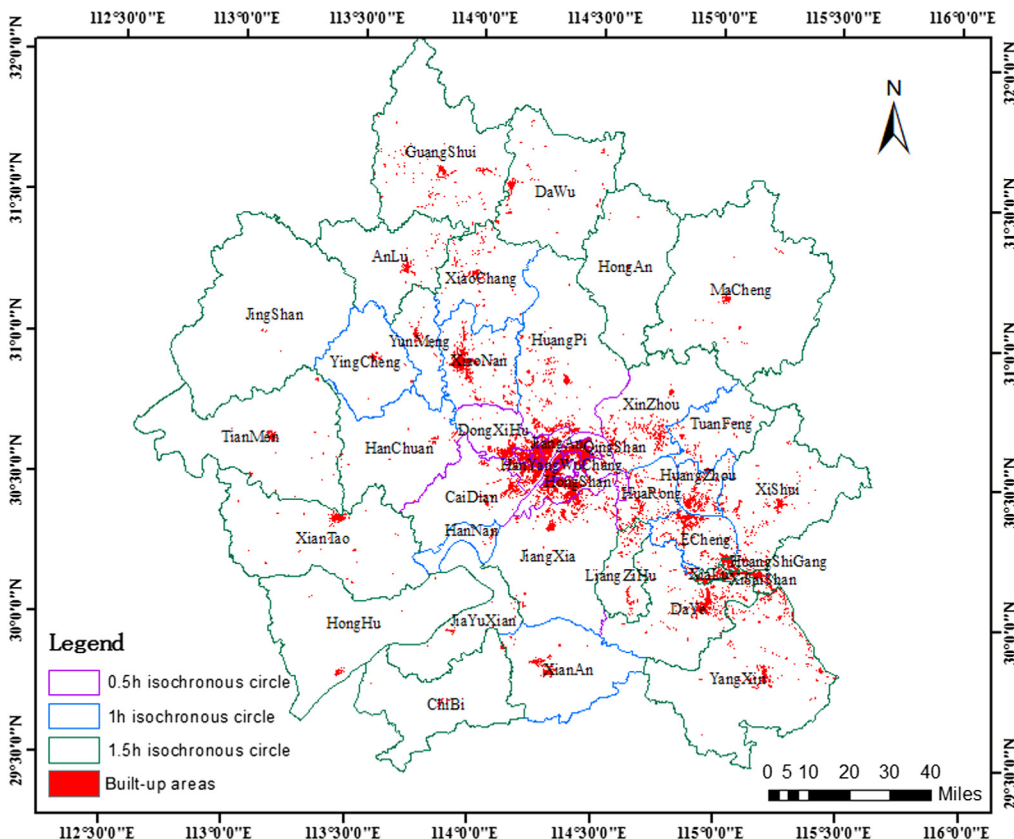

Fig. 7 Spatial distribution map of BUA.

Table 5 Comparison table of administrative district, BUA, population, and population density.

Metropolitan area	Proportion of administrative area (%)	Proportion of BUAs (%)	Proportion of permanent population (%)	Population density (people/km^2)
Core metropolitan area	13.63	44.02	32.35	1295
Subcore metropolitan area	20.41	27.49	24.28	650
Daily metropolitan area	65.95	28.49	43.37	359

administrative, about 27.49% proportion of BUA, about 24.28% proportion of permanent population, and 650 people/km^2 of population density. The largest proportion of administrative area and smallest population density appeared in daily metropolitan area. The three circles form a "target-shaped" distribution structure radiating outward from the core metropolitan area on the said four indicators.

Corridor is a linear area that spans urban and rural regions with transportation infrastructure.[40] From the perspective of regional economics, corridor is a linear system connected by transportation. It is a corridor regional economic space system formed by highly developed multimode transportation network connecting at least two or more large and medium-sized cities or urban agglomerations.[41,42] Transport lines in the metropolitan area of Wuhan such as major roads, railways, and rivers are distributed outward as Wuhan as a core. They extend northward to Xiaogan and Suizhou, spread southward to Xianan, respectively, alone with the Jingguang and Handan railways; by the Huyu expressway and Yangtze river, transport lines extend eastward to Huanggang, Huangshi, westward to Xiantao, Tianmen. Hereunder, five corridors are shown in Fig. 8. These corridors are Wuhan to Huanggang, Wuhan to Xiaogan and then to Suizhou, Wuhan to Ezhou and then to Huangshi, Wuhan to Xiantao and then to Tianmen, and Wuhan to Xianan and then to Chibi. The spatial distribution of BUA and transportation infrastructure
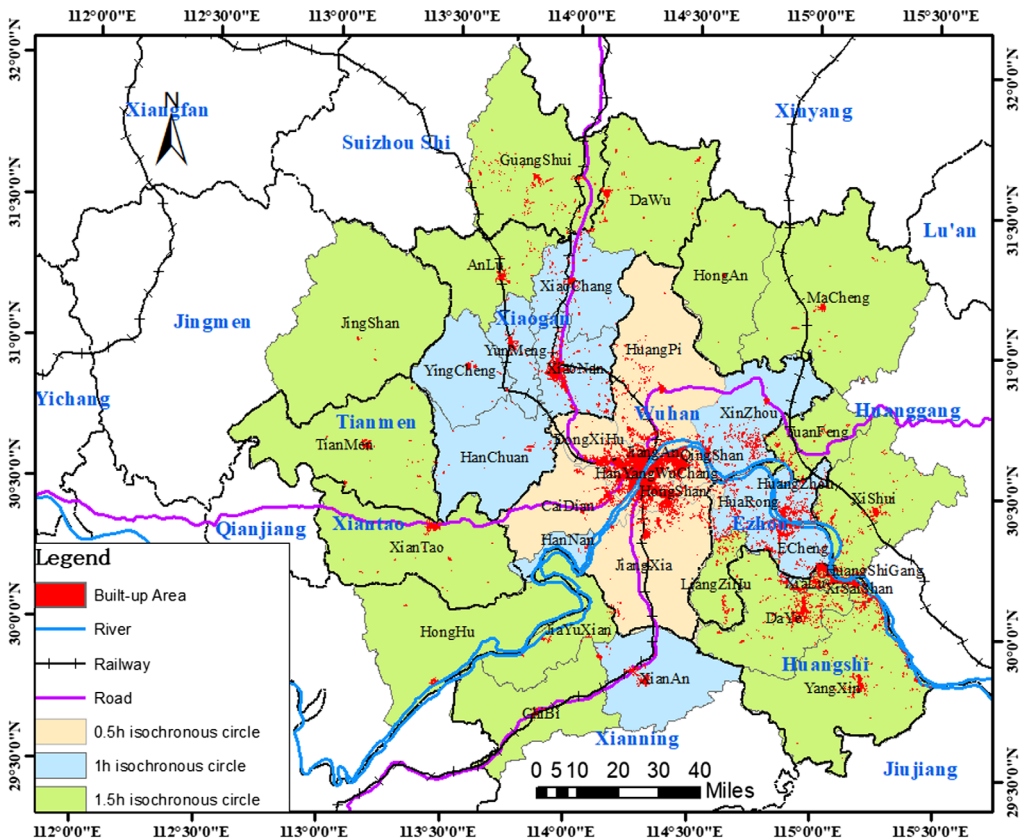
Fig. 8 Distribution map of transport lines from metropolitan area of Wuhan.

are highly consistent. Traffic corridors shorten the travel time between cities and enable people to obtain a wider space for activities and development. At the same time, they promote the sharing of various resources between cities and the optimization of resources in each functional area. They are the flow of people, material, and capital for the development of metropolitan areas and form the axis of urban development. The construction of traffic corridors will promote the process of urbanization and promote the outward expansion of BUA.

7 Conclusions

In this research, we proposed an MSDAE method for extracting BUA from GF-1 images, then we extracted BUA of the metropolitan area of Wuhan, and finally we analyzed the area of BUA, population and population density based on the extraction result and got five corridors by the distribution of BUA and transport lines. The conclusions are summarized as the following.

1. Aiming at the concentrated distribution of BUA but a large number of broken features, the MSDAE model is innovatively built by taking full advantage of the significant difference in the spatial pattern of features on different scales. MSDAE learns the features of land objects from the four scales of 1×1, 7×7, 15×15, 25×25, and identifies the types of land objects from the four scales, then determines the classification result by logistic method. By using multiscale detection windows, building multilevel feature structure, and optimizing merge rules, MSDAE can reduce the interference of similar land objects without losing the detailed information of BUA and avoid the phenomenon of salt-and-pepper noise generated by traditional pixel classification. Therefore, MSDAE effectively improves the recognition rate of BUA. MSDAE is superior to the results of single scale. Compared with other methods, MSDAE reaches high scores and provides better result than other models. Although our proposed method performs well, several issues remain to be resolved in future work. BUA has obvious feature in color space, remote sensing index, and other aspects. How to use multifeatures to integrate multiscale features to further improve the extraction accuracy of BUA and the robustness of the model needs further study.

2. Taking the metropolitan area of Wuhan as the study case, the range of the metropolitan area is 1.5 h calculated by Tencent map big data. The BUA of the Wuhan metropolitan area is extracted by MSDAE model from GF-1 WFV image. The metropolitan area of Wuhan is divided into three layers: core metropolitan area, subcore metropolitan area, and daily metropolitan. By calculating the proportion of the administrative area, BUA, permanent population, and population density, we know that the metropolitan area of Wuhan has a target-shaped distribution structure radiating outward from the core metropolitan area. Through the overlay analysis of city corridor and BUA of metropolitan area of Wuhan, it can be seen that the spatial distribution of the BUA is consistent with the spatial distribution of traffic lines; based on this, five corridors are identified: Wuhan–Huanggang, Wuhan–Xiaogan–Suizhou, Wuhan–Ezhou–Huangshi, Wuhan–Xiantao–Tianmen, and Wuhan–Xianan–Chibi. Corridor is conducive to the optimization and sharing of resources among metropolitan areas, can promote urbanization, and is of great significance to the development of metropolitan areas.

Acknowledgments

References

1. M. Taguchi and K. Narita, "The development of multi-core metropolitan areas," Vol. **23**, University of Tokyo Press, Tokyo (1986).

2. L. Wang, X. Deng, and W. Niu, "The definition and identification of urban agglomerations," *Acta Geogr. Sin.* **68**(8), 1059–1070 (2013).
3. T. Firman, "Major issues in Indonesia's urban land development," *Land Use Policy* **21**(4), 347–355 (2004).
4. C. R. Ding and X. S. Zhao, "Land market, land development and urban spatial structure in Beijing," *Land UsePolicy* **40**(9), 83–90 (2014).
5. C. G. Zhang and Y. C. Yang, "Concept analysis of metropolitan coordinating region," *City Plann. Rev.* **31**(4), 31–36 (2007).
6. X. Wang and K. Zhu, "Innovative space of metropolitan area: types, patterns and evolution: the case of Nanjing metropolitan area," *Urban Dev. Stud.* **22**(7), 8–15 (2015).
7. X. Dong and Y. Shi, "A study on the development of the metropolitan region," *Adv. Earth Sci.* **20**(10), 1067–1074 (2005).
8. J. Yuan, J. Zhou, and W. Huang, "Several long standing mistaken ideas in the theoretical studies and planning practices of Chinese metropolitan regions," *Geogr. Res.* **25**(1), 112–120 (2006).
9. T. Zakaria, "Urban transportation accessibility measures: modifications and uses," *Traffic Q.* **28**(3), 467–479 (1974).
10. C. Liu, D. Duan, and R. Yu, "Spatial evolution of space-time distance accessibility of urban-rural road network in Wuhan metropolitan area," *Econ. Geogr.* **33**(9), 698–707 (2013).
11. Z. X. Zhang et al., *Remote Sensing Monitoring of Urban Expansion in China*, p. 16, StarMap, Beijing (2006).
12. G. Yan, "Geographic area and fractal study of town's spatial distribution of Shenyang metropolitan area," *Sci. Geogr. Sin.* **36**(11), 1736–1742 (2016).
13. X. Li, "Remote sensing monitoring and simulation study of urban built-up area in Nanchang City," East China University of Technology, Shanghai (2017).
14. F. Mou et al., "Dynamic monitoring of built up area in Beijing during 1973–2005 based on multi-original remote sensed image," *Remote Sens.* **11**(2), 257–268 (2007).
15. Z. Li, X. M. Yang, and F. Meng, "The method of multi-source remote sensing synergy extraction in urban built-up area," *J. Geo-inf. Sci.* **19**(11), 1522–1528 (2017).
16. H. Lyu et al., "Long-term annual mapping of four cities on different continents by applying a deep information learning method to Landsat data," *Remote Sens.* **10**, 471 (2018).
17. G. E. Hinton, S. Osindero, and Y. W. The, "A fast learning algorithm for deep belief nets," *Neural Comput.* **18**, 1527–1554 (2006).
18. G. E. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science* **313**(5786), 504–507 (2006).
19. Y. Bengio et al., "Greedy layer-wise training of deep network," in *Adv. Neural Inf. Process. Syst.*, B. Scholkopf, J. Platt, and T. Hoffman, Eds., MIT Press, Massachusetts, pp. 153–160 (2007).
20. M. Ranzato et al., "Efficient learning of sparse representations with an energy-based model," in *Adv. Neural Inf. Process. Syst.*, B. Scholkopf, J. Platt, and T. Hoffman, Eds., MIT Press, Massachusetts, pp. 1137–1144 (2007).
21. N. Wang and D. Y. Yeung, "Learning a deep compact image representation for visual tracking," in *Adv. Neural Inf. Process. Syst.*, pp. 809–817 (2013).
22. X. Zhang et al., "Object-base land cover supervised classification for very high resolution UAV image using stacked denoising autoencoders," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **10**, 3373–3385 (2017).
23. J. Han et al., "Two-stage learning to predict human eye fixations via SDAEs," *IEEE Trans. Cybern.* **46**, 487–498 (2016).
24. T. Li, "A deep learning method for Braille recognition," in *Int. Conf. Comput. Intell. and Commun. Networks* (2015).
25. Y. Zhang, "Remote sensing image classification based on stacked denoising autoencoder," *J. Comput. Appl.* **36**(S2), 171–174, 188 (2016).
26. Y. Chen and X. Xu, "The research of underwater target recognition method based on deep learning," in *IEEE Int. Conf. Signal Process., Commun. and Comput.*, IEEE, Bilbao, Spain, pp. 1–5 (2017).

27. P.-R. Muduli, R.-R. Gunukula, and A. Mukherjee, "A deep learning approach to fetal-ECG signal reconstruction," in *Twenty Second Natl. Conf. Commun.*, IEEE, India, pp. 1–6 (2016).
28. I. Ni'mah and R. Sadikin, "Deep architectures for super-symmetric particle classification with noise labeling," in *Int. Conf. Comput., Control, Inf. and its Appl.*, IEEE, Thailand, pp. 169–174 (2016).
29. H. Tang et al., "Stacked denoising autoencoder based fault diagnosis for rotating motor," in *37th Chin. Control Conf.*, IEEE, Wuhan, pp. 5757–5762 (2018).
30. N. Xiao et al., "Adaptive feature extraction based on stacked denoising auto-encoders for asynchronous motor fault diagnosis," in *Ninth Int. Cong. Image and Signal Process., BioMed. Eng. and Inf.*, IEEE, Shandong, pp. 854–859 (2016).
31. J. Zhang et al., "Application of stack marginalised sparse denoising auto-encoder in fault diagnosis of rolling bearing," *J. Eng.* **2018**(16), 1772–1777 (2018).
32. R. Zhang et al., "Unsupervised remote sensing image segmentation based on a dual autoencoder," *J. Appl. Remote Sens.* **13**(3), 038501 (2019).
33. B. Yang and S. Li, "Multifocus image fusion and restoration with spares representation," *IEEE Trans. Instrum. Meas.* **59**(4), 884–892 (2010).
34. L. Dadao, *Regional Development and Its Spatial Structure*, p. 118, Science Press, Beijing (1997).
35. P. Vincent et al., "Extracting and composing robust features with denoising autoencoders," in *Int. Conf. Mach. Learn.*, ACM (2008).
36. W. Wang et al., "A comparative study of object tracking using CNN and SDAE," in *Int. Joint Conf. Neural Networks* (2018).
37. P. Vincent et al., "Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.* **11**(2010), 3371–3408 (2010).
38. G. E. Hinton and R. S. Zemel, "Autoencoders, minimum description length and Helmholtz free energy," in *Int. Conf. Neural Inf. Process. Syst.*, Morgan Kaufmann Publishers Inc. (1993).
39. Y. Bengio, "Learning deep architectures for AI," *Found. Trends Mach. Learn.* **2**(1), 1–127 (2009).
40. C. F. J. Whebell, "Corridors: a theory of urban systems," *Ann. Assoc. Am. Geogr.* **59**(1), 1–26 (1969).
41. Z. Liang, "Concept and research of 'corridor' in European spatial planning," *Urban Plann. Int.* **21**(1), 59–64 (2006).
42. X. Cao and X. Yan, "A review of corridor and transport corridor," *Urban Transp.* **27**(1), 50–57 (2003).

Xiaofei Mi is a PhD candidate at Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. She is an assistant professor in the environmental remote sensing application technology laboratory, Beijing, China. Her research interests include image processing and remote sensing application.

Weijia Cao received her master's and PhD degrees in computer science at University of Macau, Macao, China, in 2013 and 2017, respectively. She is currently an assistant research fellow at Aerospace Information Research Institute, Chinese Academy of Sciences. Her main research interests revolve around image encryption, image compression, and image classification. Furthermore, she is also good at other kinds of image processing, e.g., remote sensing image and medical image.

Jian Yang is an associate professor. He received his PhD in cartography and geography information system at Institute of Remote Sensing Application, CAS, China, in 2008. His main research interest is remotely sensing imagery processing and analyzing algorithms on high spatial resolution RS image, including multi-feature extraction, multisource data fusion, land cover classification and change detection of urban region. He conducts research projects on the developing of the remote sensing big data high performance platform.

Biographies of the other authors are not available.