

Deep Learning and Superpixel Feature Extraction Based on Contractive Autoencoder for Change Detection in SAR Images

Ning Lv, Chen Chen^{ID}, Senior Member, IEEE, Tie Qiu^{ID}, Senior Member, IEEE,
and Arun Kumar Sangaiah^{ID}

Abstract—Image segmentation based on superpixel is used in urban and land cover change detection for fast locating region of interest. However, the segmentation algorithms often degrade due to speckle noise in synthetic aperture radar images. In this paper, a feature learning method using a stacked contractive autoencoder (sCAE) is presented to extract the temporal change feature from superpixel with noise suppression. First, an affiliated temporal change image, which obtains temporal difference in the pixel level, are built by three different metrics. Second, the simple linear iterative clustering algorithm is used to generate superpixels, which tightly adhere to the change image boundaries for the purpose of acquiring homogeneous change samples. Third, a sCAE network is trained with the superpixel samples as input to learn the change features in semantic. Then, the encoded features by this sCAE model are binary classified to create the change result map. Finally, the proposed method is compared with methods based on principal components analysis and Markov random fields. Experiment results show that our deep learning model can separate nonlinear noise efficiently from change features and obtain better performance in change detection for synthetic aperture radar images than conventional change detection algorithms.

Index Terms—Change detection, deep learning, simple linear iterative clustering (SLIC), stacked contractive autoencoder (sCAE), synthetic aperture radar (SAR) image.

Manuscript received April 26, 2018; revised August 24, 2018; accepted September 17, 2018. Date of publication October 1, 2018; date of current version December 3, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61571338, Grant U1709218, and Grant 61672131, in part by the Key Research and Development Plan of Shaanxi province under Grant 2017ZDCXL-GY-05-01, in part by the National Key Research and Development Program of China under Grants 2016YFE0123000 and YS2017YFGH000872, in part by the Xi'an Key Laboratory of Mobile Edge Computing and Security under Grant 201805052-ZD3CG36, and in part by the 111 Project of China (B08038). Paper no. TII-18-1027. (Corresponding author: Chen Chen.)

N. Lv is with the School of Electronic Engineering, Xidian University, Xi'an 710071, China (e-mail: luvning@qq.com).

C. Chen is with the State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China (e-mail: cc2000@mail.xidian.edu.cn).

T. Qiu is with the School of Computer Science and Technology, Tianjin University, Tianjin 300350, China (e-mail: quti@ieee.org).

A. K. Sangaiah is with the School of Computing Science and Engineering, Vellore Institute of Technology, Vellore 632014, India (e-mail: arunkumarsangaiah@gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TII.2018.2873492

I. INTRODUCTION

IN REMOTE sensing change detection, it is essential to identify the changed features among the different temporal images acquired at different observation times. Synthetic aperture radar (SAR) imaging has proven to be better in change detection than other imaging sensors, especially when the ground region is observed in cloudy, rainy, or hazy weather. However, it is usually difficult to acquire clear visual features at pixel levels from a SAR image than a nature image because of speckle noise. While deep learning has been introduced recently in remote sensing for its outstanding ability to learn nonintuitive features associated with noise suppression, there are often not enough training data in change detection tasks to build the supervised deep learning network. Therefore, it is more practical to learn the change features extracted from a remote sensing image in an unsupervised way.

In the common pixel-based change detection method, the difference image is computed pixel-by-pixel by two temporal images before the change analysis with various kinds of operators, such as a log-ratio operator, a mean-ratio operator, and so on. To reduce the influence of speckle noise, Gong *et al.* [1] proposed a method based on the neighborhood-based ratio to generate the difference image. Morphological attribute profiles [2] and high-level feature representation [3] were also used to generate difference image. Guo *et al.* [4] propose a method based on a deep belief network (DBN) to obtain difference image. In difference image analysis, pixels in the difference image are divided into two groups: changed and unchanged. Thresholding and clustering approaches are the two main popular algorithms. The thresholding methods such as Otsu [5], Kittler–Illingworth minimum error threshold algorithm [6], and EM algorithm [7] are used to find the optimal threshold value to classify directly. Two-level clustering is used in [8] for classification accuracy. Markov random field (MRF) and fuzzy c-means (FCM) were combined with the proposed energy function to improve the performance of change detection in [9].

Compared with feature extraction in nature image processing, change feature extraction in a SAR image is focused on the specific change in the image content. For example, the change of season always shows the change of pixel values in the SAR image, but the true landform and geomorphic features are not changed. McNeill *et al.* [10] studied the

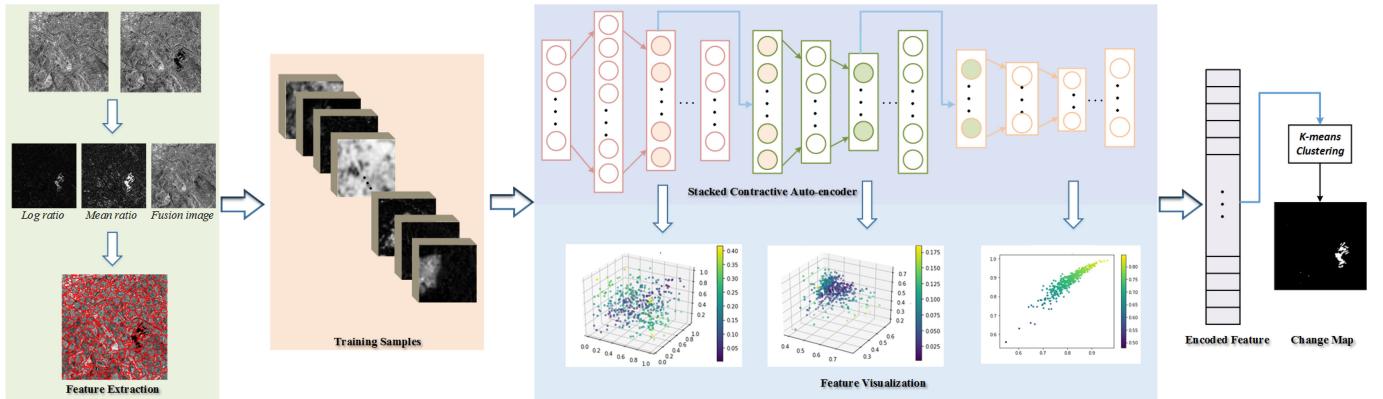


Fig. 1. Framework of temporal change feature extraction and learning.

multiseason SAR image that should focus on the changes in pixels caused by seasonal variation. In contrast, change detection in vegetated area [11] mainly detects changes in the vegetation, and that the changes in pixels caused by seasonal variation are ignored. In this case, the object-based change detection method is more effective than the pixel-based method.

In recent years, the graph-cut [12], principal components analysis (PCA) [13], and MRF [9] are used as the main algorithms for object-based feature extraction. However, PCA largely depends on the statistical characteristics of the images and is easily affected by the unbalanced data. Graph-cut and MRF algorithms have poor generalization performance with SAR images of different data sensors. Image segmentation as an object-based method is suited for acquiring the homogenous regions of objects such as temporal change. The commonly used segmentation algorithms in change detection include the mean shift segmentation algorithm [14], the improved seed-growing segmentation algorithm [15], and the fractal net evolution algorithm [16]. For the temporal images of change detection, the superpixel segmentation is more suitable [17]. Compared with an ordinary optical image, a remote sensing image has the appearance of a blurred image edge, more complex target structure, and excessive heterogeneity. These run counter to the purpose of segmentation in change detection, which is to obtain homogeneous regions that do not overlap each other. It is pointed out in [18] that since the application of traditional image segmentation methods in remote sensing images will result in oversegmentation, it is more suitable for change detection to obtain homogeneous regions by direct oversegmentation of remote sensing images. Moreover, more hidden useful features can be extracted based on oversegmentation than based on pixel.

Due to the effective object learning ability, a deep neural network has been introduced in change detection [19]. In the change detection task, the neural network can transform the original image into a high-level feature space, obtain the key information from the high-level features, and eliminate the insignificance information influence of speckle noise. The widely used neural network models for change detection include a convolution neural network (CNN) [20], [21], a DBN [22], and an

autoencoder (AE) [23], [24]. Among these models, the CNN is the most widely used, but it needs a large amount of data to train an available network, the DBN is stacked by a restricted Boltzmann machine. The appearance of the DBN solves the problem of gradient dispersion often encountered when a backpropagation algorithm is applied to deep neural networks. The AE is an unsupervised feature learning model, which learns features by minimizing reconstruction errors. A stacked denoise autoencoder is used to obtain robust high-level feature representation by the context information of each pixel of the image in [3]. A supervised contractive autoencoder is used to learn effective features from the original SAR image and difference images, respectively, in [25]. However, it is difficult to deal with limited data for change detection by the above-mentioned supervised learning method.

In this paper, an object-based change feature extraction method by deep learning is proposed to extract temporal characteristics from superpixels based on the stacked contractive autoencoder (sCAE). The flowchart is shown in Fig. 1. Superpixels are acquired by the improved simple linear iterative clustering (SLIC) algorithm. Then, robust high-level features are acquired using the unsupervised feature learning based on the sCAE. The sCAE model can be viewed as a composition of simple unsupervised CAEs, where the hidden layer of each subnetwork serves as the input layer for the next. The training samples are generated by the resized superpixels, which contain change information and geomorphic information. In Fig. 1, the feature learning process is displayed by the feature visualization of each hidden layer of the sCAE. The encoded feature is acquired, and the final change detection result is obtained using a high-level feature with k -means.

The rest of this paper is organized as follows. Section II introduces the proposed framework in detail, which contains the improved SLIC, sCAE, and the process of feature sampling and learning. The experiment results are presented in Section III to demonstrate the superiority of the proposed feature learning algorithm. In the last section, the advantage and disadvantage of the feature learning method proposed in this paper are enumerated. To facilitate reading, a notion table (see Table I) is provided.

TABLE I
SUMMARY OF NOTIONS

Notions	Descriptions
I_1, I_2	The two-input temporal images
D_1	The log ratio difference image
D_2	The mean ratio difference image
D_3	The fusion which contains geomorphic information
μ_1, μ_2	The neighborhood gray mean of I_1, I_2
l, a, b	The color vector of pixels in CIELAB color space
k	The desired number of approximately equally sized superpixels
N	The number of pixels in the image
m	The compactness of superpixels
$h_m^{(i)}$	The m th neuron in the hidden layer of the i th in the sAE
$\hat{h}_m^{(i)}$	The output of $h_m^{(i)}$
W	The weights of layers in autoencoder
b	The bias of layers in autoencoder
F_p	False alarm pixels [19]
F_n	Missing detection pixels [19]
P_{cc}	Pixel correct classified probability [19]
$Kappa$	Kappa consistency coefficient [19]

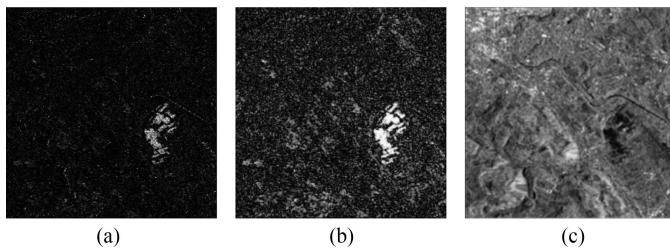


Fig. 2. Change images by the log ratio, mean ratio, and fusion image of three metrics for the Bern dataset. (a) Log ratio metric. (b) Mean ratio metric. (c) Fusion metric.

II. CHANGE FEATURE LEARNING BASED ON SUPERPIXEL

Change detection algorithms on SAR images often degrade because of speckle noise and vague imaging quality. In this section, we present a deep learning model based on the segmentation method to extract local changes in homogeneous change regions. The improved SLIC algorithm is used as the method of feature extraction to generate uniform-sized superpixels, which contain change information and geomorphic information. Then, a sCAE is modified to learn homogeneous change features from the uniform-sized superpixels to form candidate change regions. In these change regions, k -means clustering is implemented to calculate the fine change map.

A. Change Feature Metrics

Because there is only one intensity data channel for SAR image, and most of segmentation algorithms take three-color data channels as input for computation, superpixels need uniform-sized homogeneous regions and state-of-the-art boundary adherence. In order to highlight the changed regions and avoid boundary integration [26], the improved SLIC is applied. Here, a three-channel image dataset is constructed by change feature metrics using the improved segmentation algorithm.

The three metrics contain change information and geomorphic information. The difference information about the

multitemporal SAR images is obtained through log ratio and mean ratio change images. The log ratio and mean ratio are the main methods to get the difference information in change detection. The fusion image of multitemporal SAR images is used to get the geomorphic information to segment the image more accurately. The two-input temporal images are used to build the three-channel image, as shown in Fig. 2.

These three metrics are defined as

$$\begin{cases} D_1 = \left| \log \frac{I_1}{I_2} \right| \\ D_2 = 1 - \text{MIN} \left(\frac{\mu_1}{\mu_2}, \frac{\mu_2}{\mu_1} \right) \\ D_3 = \frac{1}{2}(\mu_1 + \mu_2) \end{cases} \quad (1)$$

where I_1 and I_2 are the two-input temporal images, D_1 is the log ratio difference image, D_2 is the mean ratio difference image, D_3 is the fusion image, and μ_1 and μ_2 represent the neighborhood gray mean of I_1 and I_2 , respectively.

In our framework, three change metrics are used as l , a , and b channels data in SLIC. The different weights of three metrics are set to improve the results of segmentation in practice. The CIELAB color space (the L component represents the lightness, and A and B components represent the color components) is used because it is more widely considered than the RGB and CMKY color spaces. Based on the histogram distribution of the three metrics, it is found that the histogram distribution of the log ratio and the mean ratio difference image is similar, and the fusion image is different. Thus, the log ratio and mean ratio difference images are chosen as the color components, while the fusion image as the L component. Thus, the uniform-sized superpixels are then acquired as the training samples for the sCAE.

B. Stacked Contractive Autoencoder

The AE network can extract features from the input data effectively via nonlinear mapping. In this process, a higher dimension is often set to obtain richer feature expressions. It is also necessary to regularize the expression form to eliminate useless information. For this, CAE is an effective solution.

CAE is a variant of the traditional AE that improves the robustness against noisy perturbations on the training set. A Frobenius norm of Jacobian is adopted as a regularization term. Because only the encoded feature representation is used for classification in change detection, the contractive penalty with the Jacobian term of the CAE is adopted to enhance the locally invariant and robust encoding representation, which prefers to degrade the influence of speckle noises [26]. Therefore, the entire sCAE significantly improves the discrimination and robustness of feature representation.

Specifically, for the input samples $x \in R^{d_x}$, the excitation function for feature extraction is f , and the expression of the hidden layer's feature is $h \in R^{d_h}$. As shown in the following, the penalty factor can be expressed as the sum of squares of all the dimensional differential of the feature

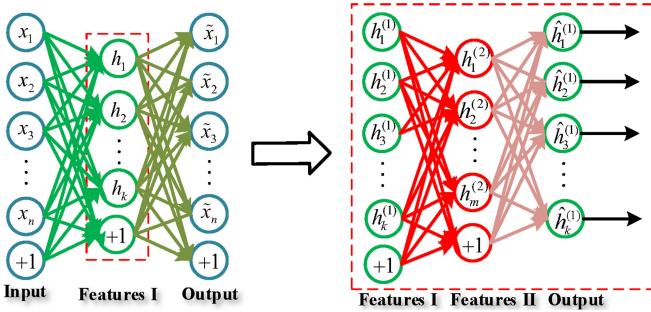


Fig. 3. Stacked contractive autoencoder.

$\|J_f(x)\|_F^2 = \sum_{ij} \left(\frac{\partial h_j(x)}{\partial x_i} \right)^2$. The cost function is as follows:

$$J = \sum_{x \in D_n} (L(x, g(f(x))) + \lambda \|J_f(x)\|_F^2). \quad (2)$$

In the cost function, the first item is the cost function of the traditional autoencoder, while the second is a regular constraint term. By adding a regular constraint, it is ensured that in the process of nonlinear mapping of training data, especially when high-dimensional feature expressions are obtained, the neighborhood information is compressed in the feature space. The points are on the high-dimensional epidemic, which means that the learned features have a local invariance.

The CAE can be stacked to build an unsupervised deep network, which has more than one hidden layer. A simple stacked CAE is shown in Fig. 3, in which the output of the first hidden layer is the input of the second hidden layer. Through the sCAE, the feature learning will get more abstract expressions from the raw data than a single CAE. With an unsupervised learning algorithm, the sCAE will not only prevent the network from overfitting when the number of labeled samples is relatively small but will also effectively provide a deep integration of features with its nonlinear mapping ability.

C. Change Feature Sampling and Learning

As previously mentioned, the remote sensing image suffers more noise than an ordinary optical image. Therefore, it is difficult to accurately locate the difference between the two temporal images. For SAR images, the speckle noise is higher, which will increase false positive detection results.

To address this, a sampling and learning framework based on SLIC and sCAE is proposed in this paper. In the framework, the improved SLIC algorithm is used to map the change metrics image from pixel feature space to superpixel feature space. Because the superpixels are going to be utilized as the training samples, they need to be transformed into patches with the same size. The choice of the patch size is essential for feature learning.

As shown in Fig. 3, a sCAE network is designed for feature learning. Initially, we use the original input sample to train the first layer of the network and get its parameters $W^{(1,1)}, W^{(1,n)}, \dots, b^{(1,1)}, \dots, b^{(1,n)}$. Then, the first layer of the network transforms the original input into a vector of hidden

element activation values (assuming the vector is A). Use A as input for the second layer and continue the training to get the second layer of parameters $W^{(2,1)}, \dots, W^{(2,2)}, b^{(2,1)}, \dots, b^{(2,2)}$. The same strategy is used for the subsequent layers. More concretely, using the sample block $x^{(k)}$ with homogenous information obtained after the segmentation as input first, the first autoencoder is trained with the original input $x^{(k)}$, and it can learn the first-order feature representation of the original input $h^{(1)(k)}$.

Then, the original data $x^{(k)}$ are input into the trained autoencoder. For each input, the corresponding first-order features can be obtained as $h^{(1)(k)}$. These first-order features are then used by another autoencoder as input to learn the second-order features.

Similarly, the first-order features are input into the newly trained autoencoder to obtain the corresponding second-order features $h^{(2)(k)}$. Next, these second-order features are used as input to learn the next feature. Based on this strategy, we trained a five-layer network structure to extract and reconstruct the change information in the difference image.

The leaky rectified linear unit function is used for nonlinear excitations on each layer in the network. This operation reduces the complex adaptability between neurons because the output of a neuron is no longer dependent on other neurons. As a result, the overfitting phenomenon is effectively avoided.

In the process of training, the dropout parameter is set as 0.5 for some layers; this operation reduces the complex adaptability between neurons because the output of a neuron is no longer dependent on other neurons. The overfitting phenomenon is effectively avoided.

During the updating for weight, root mean square prop is used for adapting the learning rate, and the Nesterov momentum formula is added. As mentioned above, the cross-entropy function is used for the cost function, and a penalty factor for regularization constraints is added.

After that, one simplified k -means clustering method is designed to loosen the condition when merged regions must be adjacent in the original algorithm. The image is divided into the changed area and two unchanged regions. After these process stages, the change map of the two temporal images is built.

III. EXPERIMENTS AND ANALYSIS

A. Introduction of Datasets

The Bern dataset [27] is a section (301×301 pixels) of two SAR images over Berns neighbor region acquired by ERS-2 and a ground truth image. The two-temporal images record the changes in the water area of Thun Lake near Bern from April to May in 1999, during which there were floods. The available ground truth (reference image) is used for the quality evaluation of experiment results.

The Mexico dataset [27] is a section (512×512 pixels) of two SAR images over the region of Mexico acquired by LANDSAT 7 SAR sensor from April, 2000 to March, 2002 and a ground truth image.

The Ottawa dataset [27] is a section (290×350 pixels) of two SAR images over the city of Ottawa acquired by the RADARSAT SAR sensor from May to August in 1997 and a ground truth image.

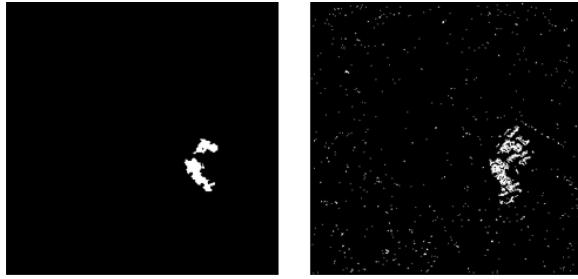


Fig. 4. Comparison between the improved SLIC and k -means. (a) Improved SLIC + k -means. (b) k -means + k -means.

TABLE II
EXPERIMENT RESULTS OBTAINED BY DIFFERENT METHODS
ON THE BERN DATASET

Algorithms	F_p	F_n	$Kappa$	Pcc
Improved SLIC	58	481	0.7115	0.9941
k -means	590	341	0.5628	0.9864

B. Results of Experiment

At first, three metrics of log ratio, mean ratio, and fusion operator were computed, respectively, as A , B , and L channels in the CIELAB space to generate superpixels. Since the log ratio and mean ratio metrics contained the change information and the fusion-metric-contained geomorphic information, change features were extracted by the improved SLIC segmentation. The data of each channel were assigned equal weights and normalized. The Bern, Ottawa, and Mexico datasets were segmented into uniformed patches. It was noted that those small isolated superpixels belonging to the true changes were always hard to classify correctly by segmentation.

In order to demonstrate the effectiveness of the change features in superpixels, the superpixels based on the improved SLIC were compared with the segmentation areas based on clustering by k -means to generate a change map. Since SLIC is similar to k -means in principle, k -means was used as the method of clustering. The number of initial clustering centers was same as the superpixels in the improved SLIC. The number was 400 in Bern datasets. The change maps are as shown in Fig. 4, and the scores of evaluation criteria are listed in Table II. Due to the full use of change information and geomorphic information, the results of superpixels are obviously better. Compared with k -means, there were less noisy FP pixels in the changed regions of the improved SLIC. In other words, the improved SLIC had the ability to suppress the inherence of speckle noise.

The superpixel maps of the improved SLIC adhered well to the boundaries of the changed area. But the inherent problem of SLIC is that it does not do well on images with complex scenes. There were still some superpixels that failed to adhere well to the boundaries of changed and unchanged area, as shown in the second column of Fig. 5. The superpixel maps on ground truth are shown in the first column of Fig. 5. The images in the second

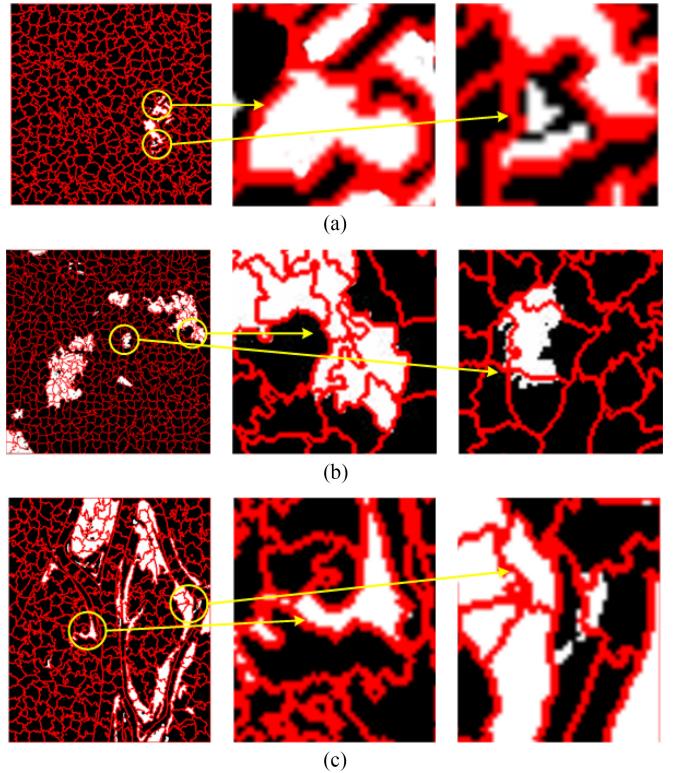


Fig. 5. Superpixel maps on ground truth of Bern, Mexico, and Ottawa datasets. (a) Bern. (b) Mexico. (c) Ottawa.

and third columns of Fig. 5 were the superpixels which adhered well to the boundaries and failed to adhere well, respectively. For change detection, the change information in superpixels that failed to adhere well to the boundaries is important.

In order to generate the learning samples that contained the neighbor information, the change feature vectors were sampled from the superpixels. These superpixels were stretched into homogeneous sample vectors in dimension n . The shape of the vectors influences the change detection result, and its effectiveness will be analyzed next. Patches that have insufficient dimensions were performed with zero padding operation, and the dimensions of the surplus were cut. From this, a set of homogenous samples containing different edge information were obtained as the training data of the network for feature learning to get the changed information of difference image.

The dimensions of input layer in the network were decided by the size of the feature vectors, which can influence the reconstruction results to a certain extent. If the size of the patches is too small, it will not only lead to a lack of homogenous geomorphic information, but it will also not meet the requirements of the input data with high dimensions for multilayer network training. If the size is too large, it will destroy the homogenous geomorphic information generated by the segmentation algorithm. Thus, in order to get the suitable size, different sizes of samples were tested. In the SLIC segmentation algorithm, the size of neighbor feature vector is as $\sqrt{k/m} \times \sqrt{k/m}$, where m is used to control the compactness of superpixels, $m = [1, 40]$, here

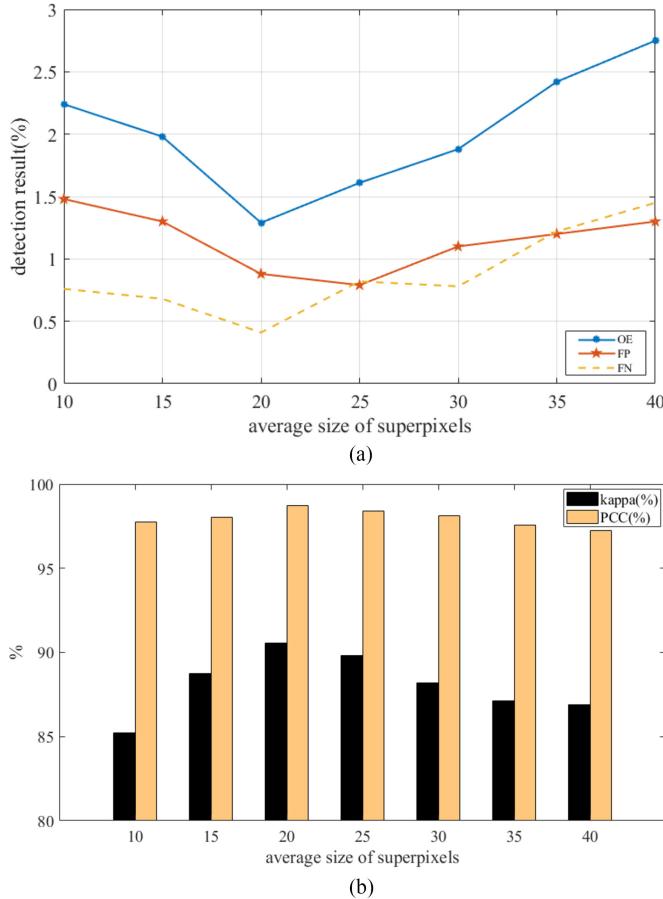


Fig. 6. Influence of a feature vector in different sizes for change detection. (a) Kappa results based on different sizes of the feature vector. (b) Error results based on different sizes of the feature vector.

$m = 20$ and k is the number of superpixels. In the experiment, $\sqrt{k/m}$ is separately set as 10, 15, ..., 40. The results of different size is shown in Fig. 6. Based on our experimental results, the size of 20 is the best choice. So, the size of homogeneous neighbor vector is set to 400.

According to the experiment results, the feature representation of the network model based on the CAE reduced the dimension of the hidden layer to get encoding features after increasing the dimension of the first layer. It can guarantee sparsity of feature sets well and reduce redundant information effectively. The CAE was first trained by homogenous samples extracted from superpixels before to reconstruct the original image. The weights of the network are used to extract as the training data to get the advanced features. Next, the original image was reconstructed from the advanced features.

As a visual example, the high-level feature of the Bern dataset is shown in Fig. 7. From the visualization results, it was found that the sCAE can not only acquire the main geomorphic information from original input, but also retain the changed information for difference analysis. The contour of the landscape and the edge of the region of change of temporal images can be clearly seen. Reconstruction based on these features suppressed the noises well to get high-quality difference images and improved the accuracy of change detection.

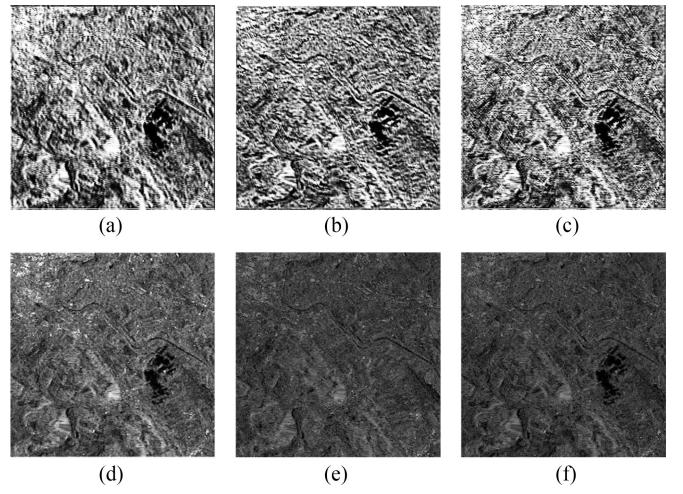


Fig. 7. Feature encoding of the hidden layer. (a) First-order feature output. (b) Second-order feature output. (c) Third-order feature output. (d) Original image of the high-order feature. (e) Reconstructed image of Bern in 1999.04. (f) Reconstructed image of Bern in 1999.05.

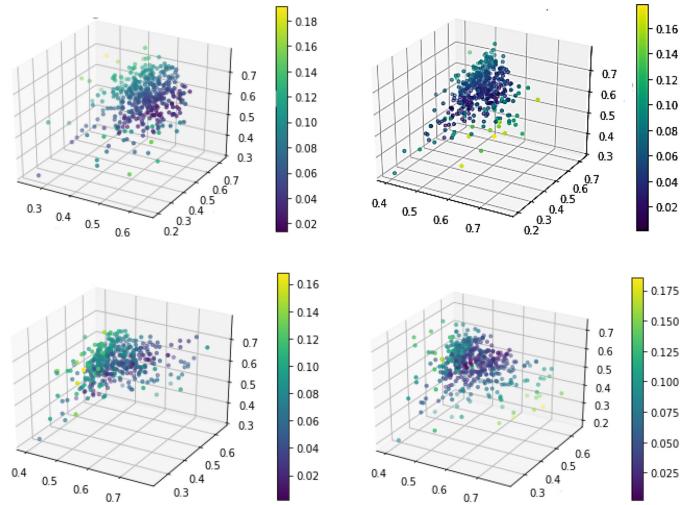


Fig. 8. High-level features in four dimensions.

Based on advanced features, the original temporal image was reconstructed well. The difference information was preserved as much as possible, especially the edge features of the changed area. The reconstruction of the original image was accomplished by using the sparsely expressed high-order features. As shown in Fig. 7, the change information was preserved, while the input samples were mapped from the sample set to the feature set with noise reduction. This means that not only the encoding change features were learned from the network, but the decoded difference image with noise suppression was also acquired based on the sCAE.

As mentioned above, the deep network was used with a 400/800/400/200/400 architecture. Three CAEs were trained for weights of the hidden layer with the greedy layerwise strategy. In order to show the relationship between the feature representations learned by the sCAE, the acquired high-level features were put into an AE to compress the input information as four-dimensional features. In order to observe the high-level

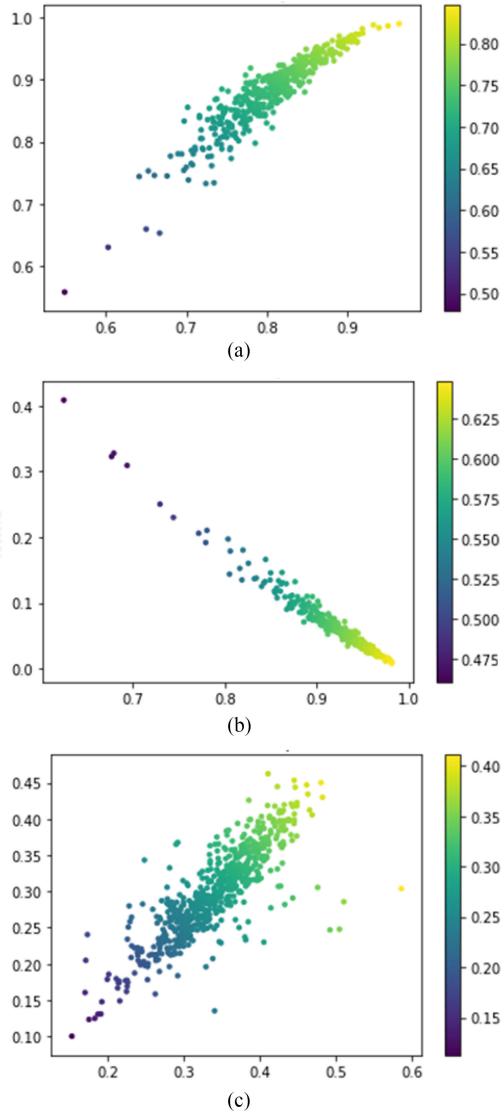


Fig. 9. High-level features in two dimensions. (a) Feature description of Bern in two dimensions. (b) Feature description of Ottawa in two dimensions. (c) Feature description of Mexico in two dimensions.

feature well, a 3-D-scatter diagram was plotted by each three-dimensional features shown in Fig. 8. The color in scatter was valued by the variance of the three vectors. It is obvious that some points were aggregated and some are dispersed. In other words, the distribution of samples matched the change information in the difference image.

In order to show the relationship more clearly, the four-dimensional feature vectors were compressed as two-dimensional vectors. A 2-D scatter diagram was plotted by the two-dimensional features shown in Fig. 9. The color in scatter was valued by the variance of the two vectors. From the scatters in Fig. 9, it is obvious that the two feature vectors were highly correlated, but there were still some outliers, which means the superpixels failed to adhere well to the boundaries. In other words, the correlation of the high-level feature representation matched the changes in superpixels.

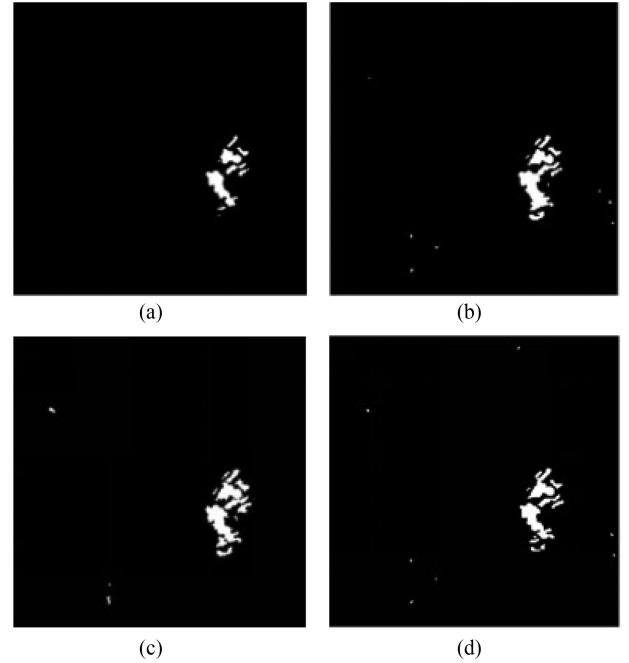


Fig. 10. Change maps by different methods. (a) PCA + k -means. (b) SRM + k -means. (c) Fuzzy + MRF. (d) Improved SLIC + sCAE.

TABLE III
PERFORMANCE COMPARISON FOUR ALGORITHMS

Algorithms	F_p	F_n	$Kappa$	Pcc
PCA+ k -means	258	146	0.8437	0.9955
SRM+ k -means	287	114	0.8363	0.9956
FCM+MRF	364	47	0.8413	0.9955
Improved SLIC+sCAE	125	135	0.9056	0.9971

The results of the four methods used in the dataset are shown in Fig. 10(a)–(d), and there are measuring performance indicators of four methods in the datasets shown in Table III. The best detection indicators are also shown in bold.

As shown in Table III, our proposed method generated better results compared with other methods, especially on the number of false alarm error and had better control on the total number of errors, making the false alarm and missing detection relatively balanced. Therefore, the Kappa coefficient also has a good performance indicator. Due to the image segmentation algorithm framework of SLIC for initial segmentation, which converts the pixel space into superpixel space, the merge area has some homogeneous change information. The sCAE learns the change regional features to allow the most similar areas to merge first, avoiding oversegmentation. The multiple change metric channel data also enhance the effectiveness of the clustering.

IV. CONCLUSION

A. Proposed Algorithm

Comparison experiments show that this method can obtain better performance of change detection than conventional PCA

and other state-of-the-art algorithms. More specifically, because the false alarm index that is usually affected by noise is improved, the Kappa index finally can show higher performance than methods based on PCA and MRF. There is a strong correlation between feature representation of fully changed superpixels and unchanged superpixels in the results of the feature learning, but weak correlation between unfully superpixels' feature representation. But we have not proposed an effective parameter to evaluate the correlation. Experiments based on the construction of multimetrics illustrates that the weights of the encoding can be learned by the sCAE to get better performance in change detection. However, determining the optimizing weights by constraining the cost function of sCAE or by deeper stacked net layers still need to be tested. The proposed method modifies the SLIC algorithm to avoid the isolated region phenomenon.

B. Future Works

With the goal of change features learning, superpixel-based change region merging method needs further study to reduce the influence of segmentation quality. The structure of the sCAE needs to be fine-tuned to improve the processing rate for extracting deeper features. The reconstruction result of the sCAE can be taken into account when generating the sample dataset with more generalization for the change feature classification.

REFERENCES

- [1] M. Gong, Y. Cao, and Q. Wu, "A neighborhood-based ratio approach for change detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 307–311, Mar. 2012.
- [2] N. Falco, M. D. Mura, F. Bovolo, J. A. Benediktsson, and L. Bruzzone, "Change detection in VHR images based on morphological attribute profiles," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 636–640, May 2013.
- [3] P. Zhang, M. Gong, L. Su, J. Liu, and Z. Li, "Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 116, pp. 24–41, 2016.
- [4] G. Cao, B. Wang, H.-C. Xavier, D. Yang, and J. Southworth, "A new difference image creation method based on deep neural networks for change detection in remote-sensing images," *Int. J. Remote Sens.*, vol. 38, no. 23, pp. 7161–7175, 2017.
- [5] L. Huang, Y. Fang, X. Zuo, and X. Yu, "Automatic change detection method of multitemporal remote sensing images based on 2D-OTSU algorithm improved by firefly algorithm," *J. Sens.*, vol. 2015, no. 3, pp. 1–8, 2015.
- [6] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 874–887, Apr. 2005.
- [7] T. Celik, "A Bayesian approach to unsupervised multiscale change detection in synthetic aperture radar images," *Signal Process.*, vol. 90, no. 5, pp. 1471–1485, 2010.
- [8] H.-C. Li, T. Celik, N. Longbotham, and W. J. Emery, "Gabor feature based unsupervised change detection of multitemporal SAR images based on two-level clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2458–2462, Dec. 2015.
- [9] M. Gong, L. Su, M. Jia, and W. V. Chen, "Fuzzy clustering with a modified MRF energy function for change detection in synthetic aperture radar images," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 1, pp. 98–109, Feb. 2014.
- [10] S. McNeill, S. Belliss, D. Pairman, and H. North, "Retrospective change detection based upon a multi-season, sparse temporal sequence of JERS-1 SAR data," in *Proc. IEEE Int. Conf. Geosci. Remote Sens. Symp.*, 2006, pp. 2353–2356.
- [11] B. Barrett, "The use of C- and L-band repeat-pass interferometric SAR coherence for soil moisture change detection in vegetated areas," *Open Remote Sens. J.*, vol. 5, no. 1, pp. 37–53, 2012.
- [12] M. Gong, M. Jia, L. Su, S. Wang, and L. Jiao, "Detecting changes of the yellow river estuary via SAR images based on a local fit-search model and kernel-induced graph cuts," *Int. J. Remote Sens.*, vol. 35, nos. 11/12, pp. 4009–4030, 2014.
- [13] O. Yousif and Y. Ban, "Improving urban change detection from multitemporal SAR images using PCA-NLM," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2032–2041, Apr. 2013.
- [14] B. Wang, S. Choi, Y. Byun, S. Lee, and J. Choi, "Object-based change detection of very high resolution satellite imagery using the cross-sharpening of multitemporal data," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1151–1155, May 2015.
- [15] L. Ma *et al.*, "Object-based change detection in urban areas: The effects of segmentation strategy, scale, and feature space on unsupervised methods," *Remote Sens.*, vol. 8, no. 9, 2016, Art. no. 761.
- [16] J. Chen, M. Deng, X. Mei, T. Chen, Q. Shao, and L. Hong, "Optimal segmentation of a high-resolution remote-sensing image guided by area and boundary," *Int. J. Remote Sens.*, vol. 35, no. 19, pp. 6914–6939, 2014.
- [17] X. Yang, W. Wu, K. Liu, P. W. Kim, A. K. Sangaiah, and G. Jeon, "Long-distance object recognition with image super resolution: A comparative study," *IEEE Access*, vol. 6, pp. 13429–13438, 2018, doi:10.1109/ACCESS.2018.2799861.
- [18] T. Wu, J. Luo, J. Fang, J. Ma, and X. Song, "Unsupervised object-based change detection via a Weibull mixture model-based binarization for high-resolution remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 63–67, Jan. 2018.
- [19] M. Gong, J. Zhao, J. Liu, Q. Miao, and L. Jiao, "Change detection in synthetic aperture radar images based on deep neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 1, pp. 125–138, Jan. 2016.
- [20] S. H. Khan, X. He, F. Porikli, and M. Bennamoun, "Forest change detection in incomplete satellite images with deep neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5407–5423, Sep. 2017.
- [21] P. Li, Z. Chen, L. T. Yang, Q. Zhang, and M. J. Deen, "Deep convolutional computation model for feature learning on big data in internet of things," *IEEE Trans. Ind. Inform.*, vol. 14, no. 2, pp. 790–798, Feb. 2018, doi:10.1109/TII.2017.2739340.
- [22] H. Zhang, M. Gong, P. Zhang, L. Su, and J. Shi, "Feature-level change detection using deep representation and feature change analysis for multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 11, pp. 1666–1670, Nov. 2016.
- [23] M. Gong, H. Yang, and P. Zhang, "Feature learning and change feature classification based on deep learning for ternary change detection in SAR images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 129, pp. 212–225, 2017.
- [24] X. Yuan, B. Huang, Y. Wang, C. Yang, and W. Gui, "Deep learning based feature representation and its application for soft sensor modeling with variable-wise weighted SAE," *IEEE Trans. Ind. Inform.*, vol. 14, no. 7, pp. 3235–3243, Jul. 2018, doi:10.1109/TII.2018.2809730.
- [25] J. Geng, H. Wang, J. Fan, and X. Ma, "Change detection of SAR images based on supervised contractive autoencoders and fuzzy clustering," in *Proc. Int. Workshop Remote Sens. Intell. Process.*, 2017, pp. 1–3.
- [26] J. Geng, H. Wang, J. Fan, and X. Ma, "Deep supervised and contractive neural network for SAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 4, pp. 2442–2459, Apr. 2017.
- [27] Y. Zheng, X. Zhang, B. Hou, and G. Liu, "Using combined difference image and k-means clustering for SAR image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 3, pp. 691–695, Mar. 2014.

Ning Lv received the B.Eng. degree in electrical engineering and automation from Xi'an Jiaotong University, Xi'an, China, in 2000, and the M.Sc. degree in circuit and system engineering from Xidian University, Xi'an, in 2004. He is currently working toward the Doctor's degree in pattern recognition and intelligent systems.





Chen Chen (M'09–SM'18) received the B.Eng., M.Sc., and Ph.D. degrees in electrical engineering and computer science (EECS) from Xidian University, Xi'an, China, in 2000, 2006, and 2008, respectively.

He was a Visiting Professor in the Department of EECS at the University of Tennessee and at the University of California, in 2013 and 2018, respectively. He is currently an Associate Professor with the Department of EECS, Xidian University. He is also the Director of the Xi'an Key Laboratory of Mobile Edge Computing and Security, and the Director of the Intelligent Transportation Research Laboratory, Xidian University. He has contributed to the development of 5 copyrighted software systems and invented more than 60 patents.

Dr. Chen serves as a General Chair, a PC Chair, a Workshop Chair, or a TPC Member of a number of conferences. He has authored/coauthored 2 books, more than 70 scientific papers in international journals and conference proceedings, such as the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, the IEEE INTERNET OF THINGS JOURNAL, the IEEE International Conference on Communications, the IEEE Vehicular Technology Conference, *Journal of network and computer applications*, *Computer Communication*, etc. He is also a Senior Member of China Computer Federation and a member of ACM and Chinese Institute of Electronics.



Tie Qiu (M'12–SM'16) received the Ph.D. degree in computer science from the Dalian University of Technology, Dalian, China, in 2012.

Prior to this position, he was an Assistant Professor in 2008 and an Associate Professor in 2013 with the School of Software, Dalian University of Technology. He was a Visiting Professor of electrical and computer engineering with Iowa State University, Ames, IA, USA, during 2014–2015. He is currently a Full Professor with the School of Computer Science and Technology, Tianjin University, Tianjin, China. He has contributed to the development of three copyrighted software systems and invented ten patents.

Dr. Qiu serves as an Associate Editor for the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, an Area Editor for the *Ad Hoc Networks*, an Associate Editor for the IEEE ACCESS JOURNAL, *Computers and Electrical Engineering*, and *Human-Centric Computing and Information Sciences*, and a Guest Editor for *Future Generation Computer Systems*. He serves as a General Chair, a Program Chair, a Workshop Chair, a Publicity Chair, a Publication Chair, or a TPC Member of a number of international conferences. He has authored/coauthored 9 books, more than 100 scientific papers in international journals and conference proceedings, such as the IEEE/ACM TRANSACTIONS ON NETWORKING, the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TCY, the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, the IEEE COMMUNICATIONS, the IEEE SYSTEMS JOURNAL, the IEEE INTERNET OF THINGS JOURNAL, the IEEE COMPUTER NETWORKS, the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETIC, the IEEE International Conference on Communications, the IEEE International Conference on Computer Supported Cooperative Work in Design, etc. There are six papers listed as ESI highly cited papers. He is a Senior Member of China Computer Federation and ACM.



Arun Kumar Sangalia received the M.E. degree in computer science and engineering from the Government College of Engineering, Tirunelveli, Anna University, Chennai, India in 2007, and the Ph.D. degree in computer science and engineering from the Vellore Institute of Technology, Vellore, India in 2014.

He is currently an Associate Professor with the School of Computer Science and Engineering, Vellore Institute of Technology. He has authored more than 100 publications in different journals and conference of national and international repute. Also, he registered one Indian patent in the area of computational intelligence. He is an Editorial Board Member/Associate Editor for various international journals. His research interests include software engineering, computational intelligence, wireless networks, bioinformatics, and embedded systems. His current research work includes global software development, wireless ad hoc and sensor networks, machine learning, cognitive networks, and advances in mobile computing and communications.