Deep Learning Semantic Statistical Features' Labeling Of SAR Image Based On K-MEANS Clustering And Wavelet Transform

Jinxi Peng*,1, Yuanqi Su1, Xiaorong Xue*,2,3, Donghong Song2,3, Xiaoyong Xue4
Aihua Wu5

1 Department of Computer Science and Technology, Xi'an JiaoTong University, Xi'an 710049,Shaanxi,China
2 School of Electronics & Information Engineering Liao University of Technology (LUT) Jizhou 121001 Liaoning ,China
3 School of Computer and Information Engineering, Anyang Normal University, Anyang, Henan 455000,China
4 The Middle School of Luonan County 726100, Shaanxi, China
5 Department of AI & software Engineering, South China institute of software Engineering, Guangzhou University, Guangzhou 510990, Guangdong, China

Email of all the authors: pengjinxi@sohu.com
yuanqisu@mail.xjtu.edu.cn
flyxxr@sina.com
598856150@qq.com
615730830@qq.com
407784919@qq.com
*Corresponding author: Jinxi Peng.; pengjinxi@sohu.com ; phone 15919690280; Xiaorong Xue ; flyxxr@sina.com; phone 13811927961.

Abstract: [Purpose] Because Synthetic Aperture Radar (SAR) image contains much speckle, at present, it's hard to obtain good results for SAR image with the traditional image classification methods. However, SAR image has rich texture information which facilitates SAR image segmentation. There is a Deep Learning method of K-means data clustering that is used to calculate the optimal k value to avoid more-segmentation or less-segmentation, the advantages of fuzzy theory and neural network are used to improve the classification accuracy of image processing.[Methods] According to the statistical characteristics of SAR images and the semantics of fuzzy neural networks analysis, an efficient image segmentation method based on Deep Learning Semantic analysis and wavelet transform is proposed to achieve precision of classification; Firstly, the texture features of SAR image are extracted by Deep Learning semantic clustering; Secondly, the SAR image are characterized according to SAR semantics. Based on one of the Deep Learning Semantic k-means algorithms, the best k-value iterations of SSE(Sum of the Squared Errors) and SC(Silhouette Coefficient) are performed and the suitable value k is selected; Finally, by Deep Learning segmentation based on the texture feature of the SAR image and the filtered gray component vector, and the SAR image is classified to facilitate the change detection of images.[Results] The experimental results show that a good k value detection is performed on the CPU/GPU platform. The SAR images is compared with two different results between the...
wavelet filtering and Deep Learning semantic method images which are single or multiple standard classified through detection-change SAR is better than the Empirical Approach, then the labels are classified for the accuracy and computational efficiency that are calculated. The classification results were accurately ameliorated. [Conclusions] The similarity calculation of k-value cluster is an important precondition, so it's necessary to select the optimal k-value suitable for the segmentation condition. The method achieves are taken an adequate SAR image analysis of effect and improvement for performance by semantic classification of k-means clustering.

1. Introduction
For the SAR can penetrate the clouds and atmosphere, the weather is changed of day and night motion is faster and faster after before, and it plays an important role in the national economy and national defense construction. Image segmentation is a key step in automatic target recognition. However, due to the presence of speckle in SAR images, it is difficult for traditional image segmentation methods to obtain good segmentation results. Therefore, it's urgent to find an efficient segmentation method to segment SAR images quickly and effectively (such as Jiao Li Cheng, Hou Biao,Yang Shuyuan (2016),Gan Lu, Wu Yan, Wang Fan, Zhang Peng, Zhang Qiang No. 2014 J. Lee and E. Pottiern 2008, Koray Kayabol and Josiane Zerubia,2013,Peter Yu, A. Qin and David A. Clausii 2012 Yaqiu Jin and Feng Xu 2013 etc), the Deep Learning is widely used by some experts, semantic classification is based on the implicit information to match the feature in the Deep Learning Emotional semantic scene, and an efficient and accurate image segmentation and classification algorithm for it.

In the field of military science and technology research and application, there are more intelligent recognition and application analysis of SAR image segmentation in the fields of land use and pollution detection, including gray value and texture feature value. It's the fact that it is difficult to Segment SAR images effectively with the information of gray values, the SAR images reflect the backscattering characteristics of radar waves from ground targets. If different ground objects have the same or similar backscattering coefficients, it's shown the same or similar gray values, which is caused confusion of them. In addition, the presence of speckle further exacerbates the image disorder. Therefore, in practical application, only gray value segmentation results are unacceptable. In the original SAR image, due to the influence of radar image speckle, the gray distribution is poor, but the texture information is very rich. Using the existing filtering methods, although the gray distribution of the image is improved, the distortion of texture information because of lost a lot of effective data. Therefore, it's combined that advantages of texture features and filtered gray values to segment SAR images and carry out prediction analysis.

Data clustering is divided the data into some aggregation classes which is according to the intrinsic properties of the data. The elements in each of the aggregation classes are the same characteristics as much as possible, and the characteristics of different aggregation classes are different, too(i.e., Cluster Analysis), it's a technique for static data analysis, and it's widely used in many research fields, not only including the machine learning data mining, pattern recognition, image analysis and intelligent analysis of biological information etc, obtaining the image features after wavelet domain image processing is usually used, while k-means clustering method to segment the image. The fuzzy neural network combines the semantic analysis of Deep Learning and the advantage of processing with specific problems, it's better segmentation and clustering performance to achieve the classification effect to analyse.

2. Analysis Of k-means Cluster
The k-means method is an iterative algorithm based on distance [1], which classifies number observation instances into value k clusters, so that each observation instance is smaller than the other cluster centers. Distance may be calculated by European distance 2-normal distance or Manhattan distance 1-norm distance) or others. Here we use European distance. To classify each observed instance to the nearest cluster center, we need to find the location of the cluster center However, it's determined the location of the cluster center, it must be known the observation examples are included
in the cluster. This seems is a question of having eggs or chickens first. In theory, this is an NP Hard problem [2], it's treated clustering analysis as an unsupervised machine learning algorithm. Different from supervised learning, such as common classification problems, the training data is tagged for classification categories, which is used to train a model to predict the classification of new data. In clustering analysis, if the data is unlabeled, to divide the similar instance into a class by the similarity degree of the eigenvalue of the instance in the data, to calculate the appropriate k value at the same time, so as not to cause the same class of data is segmented by excessive segmentation. or the segmentation is not in place, that is, the under-segmentation of different classes of data into the same kind of data analysis.

k-means consistency Segmentation by Neural Network Deep Learning Semantic region, it's is an iterative algorithm that is based on distance with the classifier of fuzzy neural network is a 3-layer feedback forward neural network.

We can do the first layer is the input layer assuming that the input sample is an n-dimensional eigenvector x and included n-neuron in the input layer, and the input layer which is the algorithm I for neurons as a vector follows number: there are three methods for calculating the central point of clustering, as follows:

1) Minkowski Distance formula : \( \lambda \) is taken values at will be what either negative, Integer, or infinite.

\[
d_y = \left( \sum_{k=1}^{n} |x_k - x_{\mu_k}|^\lambda \right)^{1/\lambda} \quad (1)
\]

2) Euclidean Distance formula: when \( \lambda = 2 \)

\[
d_y = \sqrt{\sum_{k=1}^{n} (x_k - x_{\mu_k})^2} \quad (2)
\]

3) CityBloc Distance formula: when \( \lambda = 1 \)

\[
d_y = \sum_{k=1}^{n} |x_k - x_{\mu_k}| \quad (3)
\]

The selection of the center points of the above three formulas is partly different \( \lambda \) (for the first \( \lambda \in [0,1] \)). Based on the stability method, two data subsets are generated by twice resampling of one data set, and then two data subsets are clustered by the same clustering algorithm, and two clustering results with k clusters are obtained. The similarity distribution of the two clustering results is calculated. The high similarity between the two clustering results indicates that k clusters reflect a stable clustering structure and the similarity is used to estimate the number of clusters, and the sub-method is used to test several to find an appropriate value of \( k \), as the formula 3.

4) Centroid calculation and analysis.

Method 1: take a sample and cluster with hierarchical clustering technique. K-clusters are extracted from hierarchical clustering and the centroid of these samples is used as the initial centroid. This method is usually effective, but only for the following cases: 1) the sample size is relatively small, for example, hundreds to numbers (the hierarchical clustering overhead is larger than the size of the sample). the third method selects the initial centroid and randomly selects the first point.

Method 2: take the centroid of the first pixel sample as the first point. Then, for each subsequent initial centroid, the point farthest from the selected initial centroid is selected. it's ensured that the initial centroid of choice is not only random, but also dispersed. However, this method is selected to deviate from the population sample point is solved.

5) distance measurement of centroid

The objective function is optimal or the maximum number of iterations is terminated. For the different distance measures [3-5], objective functions are often different. When Euclidean distance is used, the objective function is generally the sum of squares of the distance from the object to the center of its' cluster, as follows:
When cosine similarity is adopted, the objective function is generally the sum of cosine similarity between the object and its’ cluster centroid, as follows:

\[
\max_{i \in I} \sum_{k \in C_i} \cosine(c_i, x)
\]

In summary, the points in the assigned cluster do not get a cluster, it’s an empty Item. At this point, some strategy is needed to select a substitute centroid, otherwise, the square error will be too large. One method is to select a sample point farthest from the current center of mass. It’s eliminated the points that are the greatest influence on the total square error as candidate samples; another method is to select a substitute centroid sample from the cluster with the largest centroid sample. This will split the cluster and reduce the total number of clusters. If there are multiple empty clusters, the process of Iteration is iterative at last.

3. The k-means Clustering Algorithm Of Statistical Sample

The k-means clustering algorithm is how to find the cluster [3-5] which minimizes the trivial error criterion function. When the potential cluster shape is convex, the difference between clusters is obvious, and the cluster size is similar, the clustering results are satisfactory. Because of the linear correlation with the number of samples, the algorithm is very efficient and scalable for dealing with Big Data sets. However, in addition to determining the number of clusters \( k \) in advance and sensitivity to the initial cluster center, the algorithm often ends with a local optimum, and is sensitive to speckle and outlier information. Moreover, this method is not suitable to find clusters with non-convex shapes or clusters with very different sizes. At the time, complexity of the algorithm is \( O(ikk) \), space complexity is \( O((m+k)n) \), within them that are the numbers of iterations, \( k \) is Number of crowds, \( m \) is Number of clusters, \( n \) is dimension of sample. The computational complexity is moderate, Cluster Center is iterative method to do, Selective clustering sample as figure 1:

![Figure 1 k-means Cluster sample](image)

For k-means clustering analysis, the centroid samples are generally used for regional consistency matching. One of the common methods is to extract the characteristics of N samples at the same time, and put them together to form a N-dimensional vector space. Let it be gotten a mapping from the raw data set to the N-dimensional vector space you always need to do it at explicitly or implicitly, and then sort it out based on some kind of rule, under that rule, a same group classification has the greatest similarity within them.

3.1. Optimal k-value iterative algorithm theory for SSE (Sum of the Squared Errors), as follows:

\[
SSE = \sum_{i=1}^{\infty} \sum_{p \in C_i} |p - m_i|^2
\]

\( C_i \) is the cluster No \( i \), \( p \) is the sample point of \( C_i \), \( m_i \) is center of mass \( C_i \) (i.e., all of the mean value \( C_i \)), SSE is Clustering error of all samples, Indicates the effect of clustering.

The main thinking: with the value \( k \) increasing, The sample division will be more meticulous,
and degree of aggregation of each cluster increases gradually. Then the sum of squared errors, SSE is to changed smaller, under k smaller than the true clustering number, because of value k is fast increasing, for every extent of polymerization cluster, so SSE is rapid decline, only k is equal to cluster number, add k is the return value on which the degree of aggregation will be quickly become smaller and smaller, so the magnitude of the decline will plummet, with k The value continues to increase and tends to steady (i.e., The coordinate system between SSE and k is like a hand elbow, that k value is the real clustering number of the data), figure 2:

![figure 2](image)

3.2. SC (Silhouette Coefficient) for Contour coefficient of sample points $X_i$, as follows:

$$S = \frac{b - a}{\max(a, b)} \quad (7)$$

This case, a is $X_i$, average distance from other samples in the same cluster called cohesion degree, and b is $X_i$, average distance from all samples in the nearest cluster called resolution, the nearest cluster as following:

$$C_j = \arg\left(\min_{c_i} \frac{1}{n_{c_i}} \sum_{p \in c_i} |p - X_i|^2 \right) \quad (8)$$

p is a sample of cluster $C_j$, briefly, i.e., $X_i$ is average distance from all of the samples to a cluster as a measure of the distance from that point to the cluster, select the nearest cluster acts as a minimum value [6-8], figure 3:

![figure 3](image)

The contour coefficients of all the samples are calculated, and then the average contour coefficients are obtained. Range of mean contour coefficients $[-1, 1]$, and the closer the sample is within the cluster, the farther the sample distance among all clusters, the greater the average contour coefficient, then clustering effect is better, so maximum mean contour coefficient k is optimal clustering number.

For the center of cluster samples number $k$, $(\mu_1, \mu_2, ..., \mu_k)$ as mean-value, it's for each of component($N_1, N_2, ..., N_k$) is defined an error function defined as $S^2$ (i.e., loss function $F$), as follows:
Finding the maximum point of the convex function by iteration of gradient descent,

$$\frac{\partial F}{\partial \mu_i} = -2\times 2 \sum_{j=1}^{N} (x_j - \mu_i) \sum_{j=1}^{N} (x_j - \mu_i) \quad \text{s. t.} \quad \frac{\partial F}{\partial \mu_i} = 0; \quad (10)$$

the solution of equation:

$$\mu_i = \frac{1}{N_j} \sum_{j=1}^{N} x_j.$$  

The maximum value of the convex function is the average value of the sample points in each cluster sample, the minimum is reached sample loss function, the raw data set component \((x_1, x_2, \ldots, x_m)\), and the \(x_j\) is a \(j\)-vector of \(k\)-means cluster\([11]\), in the given number of classification groups \(k(k \leq n)\) under the conditions of value \(k\), the raw data is divided for \(k\)-classification, while variance \(S = \{S_1, S_2, \ldots, S_k\}\), \(S\) is get Minimum value:

$$\mu_i$$ is the mean value of \(S_i\).

The maximum value of the convex function is the average value of the sample points in each cluster to compute lost function \(J\):

$$J = \min \{\mu_1, \mu_2, \ldots, \mu_k\} \quad (11)$$

Minimum sample loss function, the collection of raw data is \((x_1, x_2, \ldots, x_m)\), for every \(x_i\) as \(j\) dimension vector, \(k\)-means cluster\([12, 13]\), under the condition of in which is the given number of classification groups value \(k(k \leq n)\). Divide the raw data into \(k\)-class, \(S = \{S_1, S_2, \ldots, S_k\}\), in a numerical model, the minimum value is obtained for the following expression \(S\): i.e., \(\mu_i\) gives mean value of \(S_i\).

It’s improved optimization algorithm based on clustering sample iteration is as follows:

from initial vector of random samples \(D \{\mu_1, \mu_2, \ldots, \mu_k\} \)

while \((\mu_{i+1} - \mu_i < \varepsilon)\) {

\(C_i = \phi, i = 1, 2, \ldots, k\),

for \(i = 1, 2, \ldots, m\) {

\(X = (x_1, x_2, \ldots, x_m)\), \(d = AXAX^T\), \(d_j = \|d_j - \mu_i\|, \mu_i = \frac{1}{N} \sum_{j=1}^{N} x_i \)

Computing \(k\) domain mean vector cluster of \(x_j\), \(\lambda_j = \arg \min_{i=1,2,\ldots,k} d_j\)

\(x_j \subset X\), then \(C_{\lambda j} = C_{\lambda j} \cup \{x_j\}\)

for \(i = 1, 2, \ldots, k\) {

\(\text{calculate renewal mean value} \quad \mu_i = \frac{1}{|C_i|} \sum_{j \in C_i} x_j\)

if \(\mu_i \neq \mu_{i'}\) then \(\mu_{i'}\)

else \(\mu_{i'}\) invariability } }

Output Cluster vector \(C = \{C_1, C_2, \ldots, C_k\}\) for equivalent class sample

4. Deep Learning Semantic Analysis

Image categorization is coarse-grained computing for image semantics method \([9-10]\), so, we’ll wonder design what it is converted the picture directly into a bunch of words or a paragraph of text to
describe it. After being transformed into text, we accumulate relatively deep text processing techniques through intensive learning. Firstly, it uses convolution of neural network and depth automatic encoder to extract the multi-layer features of images, and extract the implicit semantics of pictures, accordingly, an effective and accurate image search method is created by building an index image. Then make full use of a large number of Internet resources, do semantic analysis of a large number of seed images in advance, then use similar images search, according to the semantics of similar seed images to derive the semantics of new images, and use neural network hidden layer underlying feature semantics. The clustering samples, edge information, target recognition and depth learning methods are used for target classification.

The main advantage of Deep Learning lies in its automatic Learning feature model PWW[11] Suppose to design an image data, feature extraction system. \( S \); the expressing is that \( S = [S_1, S_2, ..., S_j, S_k] \), \( S \) is at \( S \) position of layer \( i \), as following figure 2 view, the inputting component \( I \) (INPUT), after system \( S \) prosing for every layer ,and the outputting is \( 0 \) (OUTPUT). Constantly adjust all parameters in each layer of the system, then correspondence with \( 0 \) and \( I \), For automatically acquire the features of the input image \( I \), then Hypothetical basis is that mean component \( \bar{I} \) through \( S_1, S_2, ..., S_n \) transform till it's no loss of image minimize information on each layer-vector. The training set of Deep semantic is computational method for this neural network till the target image of each domain is classified to an optimal value, \( S = [S_1, S_2, ..., S_j, S_k] \), it’s a different representation of the original image \( I \) that is another expressing at this case just as a mixture of testing and training sets(i.e., It’s an Enhanced learning method for hybrid for semantics neural network ), just for figure 4:

![figure 4](http://example.com/fig4.png)

**figure 4** Feature conversion of neural network model S

### 4.1. Select initialization of semantic parameters of Fuzzy Neural Network

Assume is the case that the distribution of arbitrary dimensional features in the feature space of each fuzzy class [11] is Normal distribution, which is in the center of them, and the feature vectors of different feature dimensions are independent of each other. Parameters is adjusted with Fuzzy Neural Network. \( c^j = [c^j_1, c^j_2, ..., c^j_n]^T \) represents the range of the distribution of the members of a fuzzy neural sample, for a hypothetical sample, is decomposed into total M members, each of member contains component total \( N \), then \( \sigma^j = [\sigma^j_1, \sigma^j_2, ..., \sigma^j_n]^T \) The sample that representing each component of a training sample in a paste vector. In this case, \( x^j_i \; (k=1,2,...,N_j; j=1,2,...,M) \) , then \( N_j \; (j=1,2,...,M) \) decomposed into M cluster samples, because of \( c^j \) is the center of the samples, \( j \) is the central coordinate of the member, by probability, in \( c^j = x^j \), the \( \sigma \) is defined the size of the fuzzy set of each eigenvector and variance defines the case of the reflected set and the decentralized set, obviously, there is correlation for all variance samples. When \( \sigma^j = \eta \text{cov}(x^j) = \eta E(x^j - c^j)^2 \), to select the appropriate initial value for it, those parameters with a clear physical meaning, and to set the neural network convergence is quickly solved it, in a certain case, the
classification it's needed for selecting the appropriate initial value to educe.

4.2. Select initialization of semantic parameters of Fuzzy Neural Network.

The collection of raw data is \( \{x_1, x_2, ..., x_n\} \), for every vector \( x_i \) as dimension vector \( d \) under \( k \)-means cluster data classify [13], in the given number of classification groups at condition value \( k (k \leq n) \), i.e., the groups are separated for classify \( k \) at cases \( S = \{ S_1, S_2, ..., S_k \} \), on the numerical computing model, i.e., reaching to at the minimum value \( \mu_i \) is the mean value of classify of \( S_i \) expression. Then the algorithm is following:

i. From the area \( D \) to gain random number is \( k \) as cluster for every centrality.

ii. Calculate the remaining elements separately cluster \( k \) otherness of center, these elements are grouped into clusters with the lowest degree of dissimilarity.

iii. As the result, recalculate clusters number \( k \) to each center, the method is to take the arithmetic average of each dimension of all the elements in the cluster.

iv. At all of numbers are area \( D \) that recalculate clustering according to the new center.

v. Repeat step 4, until the clustering results no change.

vi. Output the result.

Total the classify \( K \) must be as clusters numbers of \( k \), \( k \)-means to calculate the minimum value by iterative method.

4.3. Wavelet filtering

In spatial filtering, the static assumption of speckle sometimes isn’t to match the actual which is situation of the signal, so in this kind of filter, sometimes the image processing effect is not good. In fact, image analysis can not only to process. In spatial domain, also to carry out in the frequency domain. Wavelet transform has the characteristics of multi-resolution analysis. In addition, it’s very flexible in singular feature extraction and isused for time-varying filtering, but the traditional method is not available. In low Signal-to-Speckle Ratio (SNR), waveform characteristics of signal are effectively eliminated and detected. Therefore, the development of speckle suppression methods based on wavelet transform and multi-resolution analysis is a very meaningful research direction.

The lower frequency component of wavelet decomposition is decomposed with the one based on feature matching, and the higher frequency component of wavelet decomposition and the variation rule based on variance of subregion are preprocessed, and the contrast of image will decrease. In this paper, a new mixed change strategy is proposed. The low-frequency component of wavelet decomposition is changed with the rule of variation based on feature matching, and the high-frequency component of wavelet decomposition and the rule of variation based on variance of subregion are changed. B1. The strategy of low frequency component of wavelet transform Low-frequency components of wavelet filtering are fused according to the wavelet transform strategy based on feature matching rule. computing of the wavelet energy feature \((E(x, y, k))\), the mathematical formula is the following

\[
E(x, y, k) = \frac{1}{M * N} \sum_{i,j=0}^{M*N} |c(x, y, k)|^2 \quad (12)
\]

\( c(x, y, k) \) is the layer \( k \) Central pixel coordinates \((i, j)\) to match size of domain value with pixel-window size \((M * N)\), the energy is \( E(x, y, k) \). The layer at \( k \) horizontal component \( LL(k) \) is lower frequency, perpendicular component \( LH(K) \) is higher frequency. As the following figure 5
Wavelet decomposition image frequency corresponding for $F(x,y,k)$ expressing the higher and lower frequency energy components of the image at the center.

$$F_{pm}(x,y,k) = \frac{2\sum\sum C_r(x+M,y+N,k)C_M(x+M,y+N,k)}{E_r(x,y,k)E_M(x,y,k)}$$  \hspace{1cm} (13)

$F_{pm}(x,y,k)$ is at Frequency $(i,j)$ component of feature matching vector at the first level, then the horizontal and vertical components are recorded as $I_p, I_M$, Coefficient after matching feature change $C_r(x,y,k)$, corresponding value

$$C_r(x,y,k) = W_p C_r(x,y,k) + W_m C_r(x,y,k)$$  \hspace{1cm} (14)

And the covariance calculations for both components, and they change at low and high frequencies, because of the concentration of energy in the low frequency domain, but little energy in the high frequency domain. At the center of the k layer $(x,y)$ Horizontal and Vertical decomposition coefficients of pixel elements are at point $I_a(x,y,k)$ and $I_b(x,y,k)$, The covariance is

$$T_{var} = \frac{2*var_1*var_2}{var_1^2 + var_2^2}$$  \hspace{1cm} (15)

IF $T_{var} \geq T$, fetching $0.5 < T < 1$, then

$$IF(x,y) = W_1 IA(x,y) + W_2 IB(x,y), \text{ and } W_1 + W_2 = 1$$  \hspace{1cm} (16)

$W_i = \frac{\sigma_{ia}}{\sigma_{ia} + \sigma_{ib}}$ is Component of $IA, IB$ error

$$TF(x,y) = \begin{cases} TA(x,y), var_1 \geq var_2 \\ TB(x,y), var_1 < var_2 \end{cases}$$  \hspace{1cm} (17)

For each image subblock, all of them are $4m$ decomposed sub-band image. Here is for the sake of easy calculation, which is decision to carry out a single scale wavelet transform $m$ is only. It is decomposed the 4 images of the subband. Each of 4 images sub–block are absolute average energy value as a characteristic component, the energy value $e$ is:

$$e = \frac{1}{M \times N} \sum |X(i,j)|$$  \hspace{1cm} (18)

And M,N Subband X width and height. $X(i,j)$ is frequency band coefficient value, the feature vector of the sub-block are $\omega_i = [e_{i,1}, e_{i,2}, e_{i,3}, e_{i,4}], \text{ (e}_{i,k}\text{ )}$, is average energy just as subband number $k$.

SAR image characteristic of speckel is a random distribution which is a certain appearance as multiplicative property, i.e., it’s a type of broadest Gaussian distribution, as follows:

$$Z = X \cdot V$$  \hspace{1cm} (19)

Z is SAR image intensity speckel, $X$ is a random characteristics of Radar scattering from the ground targets, $V$ is stochastic process of speckel feature-vector with the logarithmic transformation of the
image multiplicative speckle converted to additive forms.

\[ \log Z = \log X + \log V \]  

(20)  

Application of Wavelet transform to variable upper bound changed equation:

\[ c_{a,b} = \int \log X \cdot \psi_{a,b}(x)dx + \int \log V \cdot \psi_{a,b}(x)dx \]  

(21)  

In this case, \( \psi_{a,b}(x) \) is a wavelet function, according to this method, the wavelet transform coefficient of speckle is determined by the numerical value of the second part, in the small wavelet domain, then speckle is relatively easy to less reduce compared to the last session.

If SAR image is transformed in semantic alogarithm[12], meanwhile, it is approximately independent of the Gabor speckle (Arsenault H. H. and April G,1976), i.e, just as Gaussian white-noise. By Fourier spectrum analysis for SAR Image analysis, the result also proved that it is reasonable. From the perspective of the energy of the wavelet transform, but Wavelet Transform of Gaussian white-noise( i.e. speckel is Gaussian-Distribution), also; in every part of the phase space is uniformly distributed because the signal is bounded. Wavelet transform coefficients are only concentrated in a small part of the phase space. In the wavelet domain, the contribution of all that wavelet coefficients in speckel modality, which the means that the energy of speckel is distributed among all wavelet coefficients, and only a small part of the wavelet coefficients is effective as the signal energy, so we set to the wavelet coefficients for two categories: one type of wavelet coefficients are transformed by speckel modality. Although, the amplitude coefficients of these kinds of wavelets are smaller variation ,then higher quantity for those, another wavelet coefficient is obtained by signal transformation, contained the result of speckel transformation which is contrary to the first condition, by the amplitude difference of wavelet coefficients to get a method of speckel reduced. The wavelet coefficient of signal to set a threshold (i.e., by the means of prior knowledge or adaptive method), Wavelet coefficients greater than thresholds belongs to the second type of coefficients, it is preserved and wavelet coefficients less than the threshold are considered is the first wavelet coefficient, it’s the speckel reduced modality and better saved the details of the image (Hua Xie, Leland E. Pierce and Fawwaz. T.Ulaby,2002),by k-means clustering image segmentation, the results of classification is ideal in the end.

The main means of wavelet filtering image tagging algorithm are as follows:

1. For main SAR image which is logarithm transformed logarithmic image;
2. With the logarithmic image is decomposed into 2 layers by wavelet transform, the absolute value of wavelet coefficients is obtained by wavelet;
   a. details the following process of sub-images;
   b. it’s divided the detail sub-image into strong and weak parts, if the absolute value of detail sub-image is greater than last one, and is treated as a strong parts, or weak parts, \( T \) is threshold value;
   c. To the percentage of \( \alpha \) multiply by sparse region to transform in wavelet domain;
   c-1. If there are multiple strong image pixels in the window, then, the central image pixel of the window held on the original value;
   c-2. If no strong component of the image pixel's window visualization, The center pixel image of the window size multiply by the percentage of \( \beta \);
3. Wavelet transform for the processed sub-image;
4. Exponential transform of Image based on Wavelet transform, to get the filtered image;
5. According to k-means clustering segmentation by texture or context, And then sorting and labeling.
5. The $k$-means Fuzzy Clustering For Main Steps Of Image Segmentation Method

To improve SAR potential for image segmentation[13], with Deep Learning Semantic segmentation method, and unite wavelet domain vector features, then comparison and Analysis of the results of classification experiments on successive Images of change detection, the SAR image the main steps of the segmentation method are as follows:

(1). After being improved algorithm, it's centered on each pixel of the original image, is obtain 8×8 window size with it, in small 8×8 images to take 4 color components within gray wavelet domain, to the 4 components as texture features of pixels.

(2). Is wavelet filtered method SAR filtering. in corresponding 8×8 small images, steps for each pixel by step (1), the average gray value is calculated as the gray feature of the pixel.

(3). In step 1 textural features and step 2 grayscale feature as feature vector of pixel, based on all the feature vectors, to transform SAR Image Segmentation into Fuzzy Neural Network, the emotional analysis of semantic features are trained much in advance, with the semantic analysis of matching Deep Learning in emotion semantics dictionary, figure out $k = 2$ segmentation, to prevent the mean value $k$ of calculation error to cause over or insufficient classifiable, so moderately calculated value $k$, in this respect, for machine learning to greater accuracy.

| Algorithm | $k$ | Accuracy : % | Time: CPU/GPU s |
|-----------|-----|--------------|-----------------|
| SSE       | 2   | 95.60        | 0.9             |
|           | 3   | 40.5         | 1.5             |
|           | 4   | 92.2         | 0.85            |
| SC        | 2   | 93.30        | 0.8             |
|           | 3   | 33.3         | 1.9             |
|           | 4   | 83.5         | 0.83            |

6. The Experimental Results and Discussions

Figure 2 is initialized SAR image, to verify the effect of the new method, to use some traditional segmentation methods to compare the segmentation of SAR images, the traditional method is based on threshold, threshold of Rayleigh distribution feature on the Edge-Information, it's detected by Robert operator information of texture etc. The fuzzy clustering $k$-means which the mean value $k = 2$ segmentation is the figure 10 and 11 are superior to figure 6,7,8 and 9, while figure 12 methods following, it's used to the segmentation results after achieving the ideal value.
Figure 6: Original gray image after change
Figure 7: Grayscale original image before change
Figure 8: Pre-change color original image

Figure 9: Changed color after original image
Figure 10: $k=2$ Pre-change detection grayscale
Figure 11: $k=2$ Post-change detection grayscale image

Figure 12: $k=2$ detection grayscale by deep analysis
Figure 13: $k=2$ Pre-change grayscale label
Figure 14: $k=2$ Post-change grayscale label

Figure 15: $k=2$ Color label image Pre-change
Figure 16: $k=2$ Post-change color label image

Statistical results of classification though prefix and post normalized change detection: figure 17 and 18
7. Conclusions
Therefore, the experimental results and the speckle SAR image segmentation is applied effectively in the field of image change detection and image retrieval with $k$-means clustering and wavelet domain image segmentation, the SSE or SC mean value is correctly selected to avoid over-segmentation or less-segmentation. At the same time, it's reduced the speckle lower, however, it's is an effect at this situation in SAR image segmentation with suitable value $k$ to avoid more-segmentation or less-segmentation for SAR image, in applications, so it's a mathematic problem NP (Nondeterministic polynomial) for the value $k$ of $k$-means method with cluster selection. To adopt wavelet filtering and Deep Learning semantic analysis with improvement of image analysis quality, then calculate the statistical eigenvalues of the samples and label them by classification. So it's an appropriate $k$ to verify further optimization and improvement in complex heterogeneous or parallel Computing improvement.

Acknowledgements
This work was supported by (1) the Science and Technology project(192102210119) supported by the Department of Science and Technology of the Fund project in Henan province, china;(2) the Fund Project of National Natural Science of China(U1204402);(2)the Foundation Project(21AT-2016-13) supported by the twenty-first century Aerospace technology Co. Ltd, china;(3)the Natural Science Research Program Project(18A520001) supported by the Department of Education in Henan Province, china;(4) the Science and Technology project(122102210462) supported by the Department of Science and Technology in Henan province, China, the Fund project for young teachers of universities in Henan province, china;(5) the college science project of South China institute of Software Engineering, Guangzhou university (ky201717); thanks a lot.

References
[1] Arsenault H. H., Aprill G. (1976) “Properties of Speckle Integrated with a Finite Aperture and Logarithmically Transformed”, Journal of the Optical Society of America, 66(11) , pp.1160-1163.
[2] Baraldi A, Parmiggian F. (1995) “An Investigation of the Texture Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters”, IEEE Transaction on Geoscience and Remote Sensing, 33(2), pp.293-303.
[3] Fang Liu, Licheng Jiao, Biao Hou, Shuyuan Yang. (2016) “POL-SAR Image Classification Based on Wishart DBN and Local Spatial Information”, IEEE Transactions on Geoscience and Remote Sensing, 54(6), pp.3250-3308.
[4] Haralick R M.(1975) “ Statistical and Structural Approaches to Texture”, Proceeding of IEEE, 67(5) , pp.780- 804.
[5] Hua Xie, Leland E. Pierce, Fawwaz T.Ulaby (2002) “SAR Speckle Reduction Using Wavelet Denoising and Markov Random Field Modeling”, IEEE Transaction on Geoscience and Remote Sensing, 40(10), pp. 2196-2212.
[6] J. S. Lee, E. Pottier (2008) Polarimetric Radar Imaging. CRC Press: Boca Raton, USA.Koray Kayabol, Josiane Zerubia (2013) “Unsupervised Amplitude and Texture Classification of
SAR Images with Multinomial Latent Model”, IEEE Transactions On Image Processing, 22(2), pp.561-572.

[7] Lu Gan, Yan Wu, Fan Wang, Peng Zhang, Qiang Zhang. (2014) “Unsupervised SAR Image Segmentation Based on Triplet Markov Fields with Graph Cuts”, IEEE Transactions on Geoscience and Remote Sensing Letters, 11(4),

[8] Peter Yu, A.K.Qin, David A.Clausi.(2012) “Unsupervised Polarimetric SAR Image Segmentation and Classification Using Region Growing with Edge Penalty”, IEEE Transactions on Geoscience and Remote Sensing, 50(4), pp.1302-1317.

[9] Soumadip Ghosh, Debasish Biswas, Sushanta Biswas, Debasree Chanda (SARkar ), Partha Pratim SARkar (2016) “Soil Classification From Large Imagery Databases Using a Neuro-Fuzzy Classifier”, Canadian Journal of Electrical and Computer Engineering, 39(4), pp.333-343.

[10] Wing-Cheong Wong, Siu-Yeung Cho, Chai Quek. (2009) “R-POPTVR: A Novel Reinforcement-Based POPTVR Fuzzy Neural Network for Pattern Classification”, IEEE Transactions on Neural Networks, 20(11), pp.1740-1755.

[11] “Machine Learning Yearning”, Author: Andrew Y. Ng 2017

[12] “Deep learning, optimization and recognition”, Author: Jiao Licheng, Press: Tsinghua University 2017 pp.241-300.

[13] “Machine learning”, Author: Tom. M Mitchell 2005 pp.125-139.