Pisco Med Publishing

# Application of BS-GEP algorithm in Remote sensing Image classification

**Lin dan ,Wu weihong**

College of information and electrical Engineering,Hebei University of Engineering,handan 056038

***Abstract:*** I It is difficult for the Traditional statistical Remote sensing classification algorithm to get higher Classifica tion accuracy under the condition of complex state. To solve this problem, BS-GEP algorithm is introduced to the study of remote Sensing image classification Problemsin this paper, to Avoid local converge NCE of the algorithm caused by the population diversity, the characteristic o f the traditional GEP, and solve the problem of getting higher classification Accuracy difficultly under the complex condition state. The experimental results have shown that classification rules based on the BS-GEP classifier can is converted into Mathema Tical expressions and obtain higher classification accuracy. Compared with GEP algorithm, the confused degreeof theclassification results are ivelyLow,and compared with maximum likelihood algorithm, the classification results are relatively clear. The classification accuracy of the classifier has been reached to.

***Keywords:*** I Remote Sensing image classification; Gene expression programming; local convergence; classification rules; Clas Sification accuracy

Remote Sensing Image classification technology has been remote sensing Digital image processing One of the important studies in the ten. with satellite remote sensing images and increasing resolution of aerial remote sensing images , Statistics based Traditional Remote sensing image classification algorithms cannot obtain higher classification accuracy , and is not good at resolving the same object Spectrum , Foreign body Spectrum , mix pixels etc title , often error points , issues such as missed points0 , so far away Sense Image Classification has been studied in the direction of intelligent classification H . In recent years , people began to study the computer pattern recognition technology should be for Remote sensing image classification , similarity to pixels into different area categories H . such as ,Paola etc BP Artificial Neural network method for classification of multispectral data , Contrast maximum likelihood experimental analysis shows that neural network classification can be

Line 0 ; He Shing, and so on MODIS , M resolution image is real data , Applying decision tree algorithms to agricultural vegetation classification issues under 0 ; Liuxiaoping, etc, proposes a new algorithm based on ant colony Intelligent Rule Digging Tunneling New classification algorithm , simulating ants ' way to find food construct classification rules 7. These algorithms, although to a certain extent Improve the precision of remote sensing image classification , But there are not enough . Neural Network algorithm long time required for training network , indeed The best number of layers and the number of neurons in each layer there are _ set Sleepy hard 8 ° ; Expert system approach to the complexity and type of remote sensing Spectra recognition rule inaccuracy ; classification based on decision tree methods Problems with dependencies between datasets in the classification process and past degree fitting M. These deficiencies make them end up with no

can be widely applied .

A gene expression programming algorithm based on Group Policy BS -



GEP ( Expression programming based on block strategy ) applies to remote image classification problems , build on BS-GEP Classifier model , by adjusting genetic operator extraction points class Rule . uses this classifier model to classify satellite remote sensing data as "rich information for the " , to make the classification result heightquantization , area statistics and land use .

**1. BS ~ GEP Algorithm Overview**

2001 Year Ferreira first proposed gene expression programming calculation

Method GEP (Gene Exp ression Programming ) like . This is

_ New Data mining technology . the algorithm borrows from biological genetic gene expression patterns , inherits and extends the genetic algorithm GA and LegacyBiography design GP Evolution idea , differs in individual encoding and result performance aspects , to slow _ Solutions GA and GP Computational complexity and difficulty creating new The possibility of change M. is now widely used in the areas of classification clustering , rule Mining and Time series analysis, and so on M.

As a result of genetic manipulation of gene expression programming is a legal Chromosome is a random combination of , more Adaptive Individuals more easily multiply . when multiplying to one fixed algebra , The Enter of the population will gradually appear to the trend of high adaptive individuals. . This trend increases the average fitness level for a population , destroys species Group Diversity , cause evolutionary error in direction , make it evolve go to an optimal branch , lose global best cancan 14 5. This local convergence in the evolutionary algorithm of the population multiplication "" is very likely to perpetuate the evolution of its species .

For reference to a The idea of a-like evolutionary algorithm to preserve the diversity of the population , introducing a group based policy BS- GE algorithm , through dynamic adjustment Its genetic operator , Avoid traditional genetic algorithms falling into local convergence . 1.1 BS-IEP scenario Design

1) set /,(i = 1,2,..., n) to individual ' Adaptive degree , Press / , Values sort the individual in order from small to large , Converts the population To m Group with B. (J = 1,2,,... , m ) Indicates the number of individuals mesh allow less N / m ]),Bj The maximum fitness value small on B. + 1 Minimum fitness value in /(J+1) -min, / / -max <

F (J+1) - min .

2) because the individual fitness value in each interval is close to, mining scales the fitness function with a linear or power function transformation change ; and then select individuals from each interval to take advantage of the Web Roulette Select Optional genetic operation .

3) Sets the genetic operator for on the basis of the differences between individual distinctions of each interval . The mutation is the main factor that affects the diversity of the population factors So in order to maintain higher diversity in the population , on all The improved algorithm in the genetic operator only adapts the mutation operatorwhole . set a smaller mutation for an individual's superior range rate , An individual's poor interval sets a larger mutation rate .

1.2 Fitness function Design

in the symbolic regression problem of gene expression programming , Ferreira Two comments based on absolute error and relative error the price model itself has drawbacks . so , in statistics , to comment Price Two sets of data compliance methods more of a decision factor _r2 (coefficient To Determination ), calculation formula R2 =1 -sse/sst. (1 )

N

where , SSE = Z (y, - y,) is the actual observations with each back

I = 1

residuals squared and ; SST = Z (y , - y) 2 Yes SSE and

I = 1

SSR ( regression squared ) Z (y , -y) 2) Total ; y , is the actual

I = 1

Observations ; y is the average value of observations ; y , is the regression of the observation value value . then , to design the fitness function as

F = n x /x [ R] [*] 2. (2)

where , is the number of samples . because of SSE<SST , have 0<r 2 < 1, To know that , The range of adaptive values is (0 , x 100). when an Individual fitness value in a comparison set in the area is compared to , Population Fitness value is difficult to improve , to avoid overall evolutionary efficiency affected , add / multiply factor to linear magnification .

1.3 improvements to the mutation operator

Dynamic Variation of the mutation rate , ensures that the improved algorithm has has adaptive . Design variance probability function to

F ,+ Magic max . ,

P un , t = PM "" (+) () () * * * * X e f T-c . (3)

where , Pm , represents A T Generation , The mutation rate for the interval ; Pm is a Constants , set before multiplication , takes a general value in (0,

0) between ;F , indicates current interval average fitness ;( F,) Max

is the maximum of the average fitness for each interval ; ft is T for a-generation population Average fitness ; C for maximum individual fitness during reproduction

N x.100.

Is provided by the type (3) Know , for each range of P M with F , increases, minus small , increasing evolution algebra ( F ,) Max Increase ) and Decrease ; P m is scoped to 0, IpM ],_ in the standard GEP Variation of

E

Rate 0.03 ~ 0.05 Range Floating .

**2. Is based on the BS ~ GEP Remote Sensing Image classification model for**

2. 1 chromosome encoding

in the Genetic evolution algorithm , Each in the population ' chromosome All represent a problem-solving one resolution , This allows you to associate a non-numericThe problem translates to numeric problems . corresponds to actual application , That is, each chromosome in an individual can be encoded and transformed into a Remote Sensing image classification rule . to encode a specific application into a legacy pass algorithm recognizable chromosome encoding is the key to implementing the algorithmis located in .

The chromosome encoding specific scheme is as follows . non-terminal collection T Select with TM Medium Band 2,3,4,5 and 7 this 5 band , is a total of 5

variable : T = {B2, b3,b4, B 5, B 7} ; Function Collection Select Common functions :F ={+,-, X ,/}; because of the complexity of the problem , gene number is 2; head length is H=5 Select {+} is the base The connection operator is . computes each gene tail length t = 5 X (5-1) + 1 =, Each gene has a total length of L = h +t = num; The connection operator is { + } , So the total chromosome gene length to then have table 1 two gene strings shown in .

2.2 Genetic operator parameter settings

The genetic operator varies according to the degree to which an individual adapts to the environment Genetic Operation , to achieve the colony's fittest evolutionprocedure . its 3 basic operator is : Select , cross , Variant . Select Optional combination Web Roulette " " and "" Elite policy " law " Chiki This idea is based on the fitness value of each chromosome in a population The probability that it is copied to the next generation . the crossover is to replace the two-parent chromosomes with the reorganization ; Variation refers to the following from the parent tree Machine Select mutation to generate a subtree instead of the child on the node

Tree .

**3. Experimental results and analysis**

with TM Medium Band 2,3,4,5 and 7 this 5 A-band Lake North Zhang River watershed image for experimental data , Spatial resolution is m, Image After sample is 651Pixel x The Pixel . meter-like points area mostly covered with woodland and arable land except body of water , resident and paddy fields are very few . so that objects are divided into 6 class : body of water, forest to , residents ' place , Paddy Field , arable land , Shadow . the grayscale values for each of the partial sampling points for eachobject are as table 3 shows .

The mean value is used for the gray value of the sampling points in the same band. count , final figure 2 The spectral diagram shown in is .

shades of pixels corresponding to the pixel DN value , to obtain the The return value of the previous chromosome . compares the return value with the threshold ,sentence Break the pixels category area .

Random sampling selection points as training data , through the Genetic iteration calculation for generation . working with training data samples of correct classification465 ,, Overall classification precision reached 93%. Training Results and actual classification results indicate water classification effect fruit better ; poor farmland and woodland classification effect , is more confusing than large ; Other objects are also partially confused . bs~gep algorithm category Kappa coefficient is 0. , GEP for algorithm classification Kappa coefficient = 0. 85. due to GEP local convergence , causes forest and tillage to be confusing to a greater extent . bs~gep Test Precision table 4 The .

under the same hardware conditions , and using the maximum likelihood classification The results of the are compared to the following .

1) training time for a training sample for to 6 s, and the maximum likelihood classification requires Ten s . This indicates that the gene expression programming algorithm that changes is more efficient in processing than statistics Learn more advantages .

2) pass to 6 class target objects Training , BS~GEP algorithm The gene expression classifier model defined by the is divided into the entire image class results as shown 3 : , The classification result of the maximum likelihood classification ( diagram 4 shown .

from Diagram 3 and Diagram 4 Looks at the whole , Is based on the BS ~ GEP The is divided into The class model works better for object category classification .

**4. closing**

The main issues involved in the model classifier built by include the base due to encoding , Initial group settings , fitness function Design , Genetic The design of the operation and the settings for the basic parameters , These issues are related to the The evolution of the model and its classification results have a significant impact on . the real result shows that : build on BS~GEP Remote sensing image for algorithm classifier Easy to implement, the classification rules extracted by can be converted to straightmathematical expression form , to improve the environment around the sampling point Classification Precision in complex , to enter _ Step Research provides a reference to .

# Reference


1. Huangning , Liu Xiaojun , Zhu Minhui . Remote Sensing Image classification technology research J . North China works College Journal of Test Technology , 2001,(2) : 86~92.
2. Wang Yida, Shen Xiling, Xie jiongguang. . Overview of classification methods for remote sensing images J ]. Remote Sensing letter 5 , 006 () : 6740-
3. Ondeyun , Yang li. . Application of artificial intelligence technology in remote sensing image classification J . computer emulation , at the (6) : 240~243.
4. Yang . Classification of multispectral Remote sensing images based on pattern recognition method [D]]. Changchun : Northeast Normal University , 2011:641.
5. PAOLA J D, Schowengerdt RA . A Detailed Comparison of back-propagation Neural network and maximum-likeli

hood classifiers for Urban Land use Classification J . IEEETransactions on Geosciences Remote Sensing , 1995 (4): 981996.
6. He Shing, Dong Zhihai. , . Remote Sensing Vegetation classification technology based on decision tree system J ]. Surveying and mapping science , 2013,39 (1): 83-86.
7. Liuxiaoping , Licha . A new classification method for remote sensing images based on ant colony Intelligence J ]. Journal of Remote Sensing , 2008,12 (2): 253 262-
8. Lu Jingguo. . Remote Sensing image classification and modeling based on neural network integration J . Mapping Bulletin , 014 (3): 1740.
9. Zengjianhang , Wei Mong, Wang Huihui, etc. . Fuzzy Classification of remote sensing images based on knowledge method J . Journal of Surveying and Mapping science and technology . 2008,25 (3): 172475.
10. Baishulian , Bajars , Hasky Chige . Is based on the C 5.0 Remote Sensing Image decision tree for Classification Experimental study J . Remote sensing technology and Applications , 014,9 (2): 338-343.
11. FERREIRA C. Gene Expression programming : A Adap tive algorithm for solving Problems J] . Complex Syste ms, 2001,13 (2): 87429.
12. Liu has a , Lu Yinan, Wang fei , etc . Overview of Genetic programming methods J ] count computer Research and development , 2001,38 (2): 213422.
13 . Peng , Yuanchang , Maichon . A review of the theory of gene expression programming J . Computer Research , 2011,28 (2): 413-419.
ZHANG Y Q, XIAO J , SUN S J. . bs-gep algorithm for Pre diction to Software failure Series J] . Journal to Software , 2012,7 (1): 243-248.
15. Peng , Yuanchang , Maichon , , and so on . A theoretical study of gene expression programming Statement J . Computer Research , 2011,28 (2): 413-419.