

REMOTE SENSING IMAGE CLASSIFICATION OF GEOEYE-1 HIGH-RESOLUTION SATELLITE

YANG BOGANG ^{a,b}, XIN YU ^{a,b*}

^a Beijing Institute of Surveying and Mapping, Beijing, 100038, China

^b Beijing Key Laboratory of Urban Spatial Information Engineering, Beijing, 100038, China, china_yuxin@163.com

Commission VI

KEY WORDS: Texture Classification, Pattern Recognition, High-resolution, Satellite Image

ABSTRACT:

Networks play the role of a high-level language, as is seen in Artificial Intelligence and statistics, because networks are used to build complex model from simple components. These years, Bayesian Networks, one of probabilistic networks, are a powerful data mining technique for handling uncertainty in complex domains. In this paper, we apply Bayesian Networks Augmented Naive Bayes (BAN) to texture classification of High-resolution satellite images and put up a new method to construct the network topology structure in terms of training accuracy based on the training samples. In the experiment, we choose GeoEye-1 satellite images. Experimental results demonstrate BAN outperform than NBC in the overall classification accuracy. Although it is time consuming, it will be an attractive and effective method in the future.

1. INTRODUCTION

Image classification will still be a long way in the future, although it has gone almost half a century. In fact, researchers have gained many fruits in the image classification domain, but there is still a long distance between theory and practice. Therefore, image classification is an interest, open area and a bottle problem for photogrammetry and remote sensing. However, some new methods in the artificial intelligence domain will be absorbed into the image classification domain and draw on the strength of each to offset the weakness of the other, which will open up a new prospect. Consequently, this paper will apply a new method in the artificial intelligence domain---Bayesian networks (Friedman, N., 1997), to image classification domain. In general, Bayesian networks represent the joint probability distribution and domain (or expert) knowledge in a compact way and provide a comprehensive method of representing relationships and influences among nodes (or feature variables) with a graphical diagram. Accordingly, by virtue of advantages of Bayesian networks we will try our best to explore a new road to texture classification of High-resolution satellite images for achieving the automatization and intelligentization of photogrammetry and remote sensing.

Since 1988, Pearl et al. had put up the concept of Bayesian Networks, which is a powerful of inference under conditions of uncertainty. However, In the beginning, Bayesian Networks were not considered as classifiers until the discovery that Naive Bayesian Network, a very simple kind of Bayesian Networks that assumes the features are independent given the class attribute (node), are surprisingly effective (Langley, P ,1992). From then, some researchers started to explore more deeply into Bayesian Networks as classifiers. Actually, the “Naive” independent assumption in Naive Bayesian Network can not be hold in many cases, so some researchers wondered whether the performance will become better if we relax the “strong and unrealistic” independent assumption among features (Yu Xin, 2005; D.Heckerman, 1995). Thus, this paper puts up a new

method, to construct the topology structure of Bayesian Network Augmented Naive Bayes (BAN), and it can resolve the forenamed problem (or assumption), because it allows arbitrary relation (arc) among features, which can be obtained in terms of training accuracy based on training data (samples). In addition, in order to validate the feasibility and effectivity of BAN, we will apply BAN to the texture classification of High-resolution satellite images.

This paper is organized as follows. In section 2, we review some basic concepts of Bayesian Networks Augmented Naive Bayes (BAN) and then we introduce the mathematic model of BAN in detail in section 3. Then in section 4 we test on High-resolution satellite images (GeoEye-1) based on BAN. Finally Section 5 describes the experiments and draws some conclusions.

2. BAYESIAN NETWORKS FOR TEXTURE CLASSIFICATION OF HIGH-RESOLUTION SATELLITE IMAGES

In this section, we simply introduce some basic concepts about Bayesian Networks and then apply it to texture classification of High-resolution satellite images.

2.1 Bayesian Networks

Bayesian Network is one kind of effective inference methods in artificial intelligence and expert systems. In the Bayesian networks, the nodes express the variables and the arcs express a probabilistic relationship among the connected variables. In the Figure 1 we can notice that Bayesian Network topology structure is a directed acyclic graph, provides a graphical view of variables' relationships and Bayesian Network parameters specify how a node depends on its parents with quantitative probability. If the node $X_i \rightarrow X_j$, it implies that the node

X_i has direct influence upon the node X_j . Further more, the

* Corresponding author.

node X_i is the parent node of the node X_j and by contraries the node X_j is the child node of the node X_i , as a rule we denote the parent set of the node X_i by $Pa(X_i)$. Accordingly, we can compute the joint probability of all the variables based on the definition of Bayesian Networks.

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i)) \quad (1)$$

In fact, Bayesian Network is accustomed to naming Bayesian Networks Classifiers in the classification domain (Cheng, J., 1999). To relax the independent condition in Naive Bayes Classifiers (NBC), some researchers put up Tree Augmented Naive Bayes Classifiers (TAN), which extends Naive Bayes by allowing the nodes to form the topology structure similar to a tree. Whereas, the tree topology structure can not express inherent relations among features either. Wherefore, this paper will apply Bayesian Network Augmented Naive Bayes (BAN) to classification and BAN classifiers that extend Tree Augmented Naive Bayes (TAN) classifiers by allowing the features to form an arbitrary graph rather than just a tree (Friedman, N., 1997).

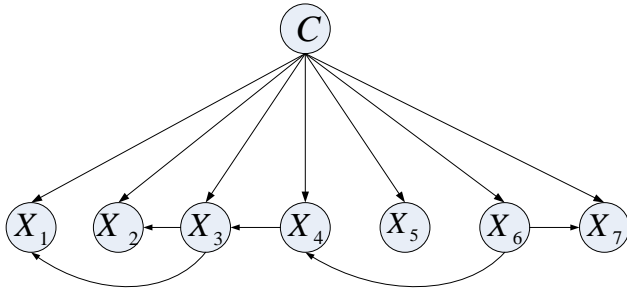


Figure 1. An Example of BAN Applied In the Classification

Figure 1 is an example of BAN applied in the classification domain. The node C denotes the class label and X_1, \dots, X_6 and X_7 denote the texture feature that are extracted from image classification unit. Thus, according to the above definition of Bayesian networks, we can compute the following probabilities.

$$P_a(C) = \emptyset, P_a(X_1) = \{C, X_3\}, P_a(X_4) = \{C, X_6\}.$$

Bayesian Networks learning includes two steps: topology structure learning and parameters learning (Cheng, J., 2001). In fact, learning the topology structure means to get some relationship among the features and the parameters learning means to estimate the parameters of the assumed probability density (distribution) from the training samples with known class label. Whereas, structure learning is more difficult than parameters learning (D.Heckerman, 1997), which is still an open problem since Bayesian Networks are put up. Hence, in order to get a Bayesian Network from data, some researchers put up many methods. There are two kinds of methods to learn the topology structure. One is the scoring-based learning algorithm (Jiebo Luo, 2005), that find one certain structure that maximizes the Bayesian, MDL or Kullback-Leiber (KL) entropy scoring function (D.Heckerman, 1995) and the other is CI-based algorithm (Yu Xin, 2007) (the conditional independent test such as Chi-squared test and mutual information test).

In this paper, we put up a new method to acquire the topology structure of BAN in terms of the training accuracy based on

training samples. Firstly one certain topology structure is given as initial structure and then estimate the relevant parameters based on training samples. And we regard the training samples as test samples to test and get the overall accuracy of the training samples, which is named the training accuracy. Then search all possible network topology structures and get the corresponding training accuracy. Among them, there exists the maximum training accuracy, whose topology structure is considered as the best one for fitting the training samples (YU XIN, 2008).

2.2 Mathematic Model and Inference of BAN

Suppose X_s is one arbitrary feature in some BAN and for sake of simplicity the parent set of X_s is denoted as X_p (i.e. $Pa(X_s)$). Ordinarily, we use a capital letter, like X to denote a random variable and a particular value of a random variable will be denoted with a lower case letter, in this case x . Commonly, it is reasonable to assume that X_s and X_p follow multi-normal distribution. In addition, X_s and X_p can constitute a new $(1+n)$ -dimensional normal random vector X with corresponding mean vector μ_X and covariance matrix D_X .

$$X_s \sim N(\mu_s, D_{ss}), X_p \sim N(\mu_p, D_{pp}) \quad (2)$$

$$X^T_{(n+1) \times 1} = \begin{bmatrix} X_s & X_p \end{bmatrix}_{1 \times 1 \quad n \times 1} \quad (3)$$

$$\mu_X = \begin{bmatrix} \mu_s \\ \mu_p \end{bmatrix} = \bar{X},$$

$$D_X = \begin{bmatrix} D_{ss} & D_{sp} \\ D_{ps} & D_{pp} \end{bmatrix} = \frac{1}{n} (X - \mu_X)(X - \mu_X)^T \quad (4)$$

The joint-normal probability density can be specified in following ways (Cui, Xizhang, 2001).

$$f(X) = (2\pi)^{\frac{n+1}{2}} \times |D_X|^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2} \begin{bmatrix} X_s - \mu_s \\ X_p - \mu_p \end{bmatrix}^T D_X^{-1} \begin{bmatrix} X_s - \mu_s \\ X_p - \mu_p \end{bmatrix}\right\} \quad (5)$$

According to Bayes' rule, we can compute the conditional density of X_s given $X_p = x_p$

$$f(X_s | x_p) = (2\pi)^{\frac{n+1}{2}} \times |\tilde{D}_{ss}|^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2} \begin{bmatrix} X_s - \tilde{\mu}_s \end{bmatrix}^T \tilde{D}_{ss}^{-1} \begin{bmatrix} X_s - \tilde{\mu}_s \end{bmatrix}\right\} \quad (6)$$

Where $\tilde{\mu}_s$ and \tilde{D}_{ss} denote the conditional mean and covariance of X_s given $X_p = x_p$ respectively.

$$E(X_s | x_p) = \tilde{\mu}_s = \mu_s + D_{sp} D_{pp}^{-1} (x_p - \mu_p) \quad (7)$$

$$D(X_s | x_p) = \tilde{D}_{ss} = D_{ss} - D_{sp} D_{pp}^{-1} D_{ps} \quad (8)$$

Then the conditional probability of X_s given $X_p = x_p$ can be got by

$$P(X_s | x_p) = f(X_s | x_p) \times dX_s \quad (9)$$

Where dX_s denotes extremely small step and is regarded as one constant in the computation.

Thus, based on formula (1) and (9), we can get the joint probability of all variables based on the following formula.

$$P(X_1, X_2, \dots, X_n, C_i) = P(C_i) \prod_{s=1}^n P(X_s | X_p) \quad (10)$$

Where C_i denotes the class label (variable) and $i = 1, \dots, m$ represents the relevant class.

In terms of the conditional probability formula, we can have

$$P(C_i | X_1, \dots, X_n) = \frac{P(X_1, \dots, X_n | C_i) P(C_i)}{P(X_1, \dots, X_n)} = \frac{P(X_1, \dots, X_n, C_i)}{P(X_1, \dots, X_n)} \quad (11)$$

As $P(X_1, X_2, \dots, X_n)$ is a constant and is independent of C_i , so we have

$$P(C_i | X_1, X_2, \dots, X_n) \propto P(X_1, X_2, \dots, X_n, C_i) \quad (12)$$

In the final, when posterior probability $P(C_i | X_1, X_2, \dots, X_n)$ is maximal, we can decide the class C^* will be $\max_i \{P(C_i | X_1, X_2, \dots, X_n)\}$.

2.3 Texture Extraction and Description

Texture features are very important in image classification domain for a long time. So far, many approaches have been proposed. Usually, these methods can be divided into two types: structural (transform-based) texture features, such as Skewness statistics (X_1), information entropy (X_2), and inverse difference moment based on the gray co-occurrence matrix (X_3), and statistical texture features, such as . the mean of LL sub-image (X_4), and standard deviation of LH sub-image (X_5) and HL sub-image (X_6) at the first decomposition level through the Symlets wavelet transform (Yang S, 2002) and fractal feature (X_7).

2.4 The Classification Scheme

To show BAN applied in the texture classification of High-resolution satellite images in detail, the complete classification scheme is summarized below (YU XIN, 2008).

- ① Training and testing samples of each class are chosen from the whole image;
- ② Extract seven kinds of texture features from each classification unit;
- ③ Arbitrarily select some network topology structure as the initial structure and estimate the parameters $\tilde{\mu}_c$ and \tilde{D}_{cc} of each class (in the formula (6)) based on the training samples by the formula (7) and (8). And then the training samples are regarded as the testing samples to be classified and the initial training accuracy is obtained.

- ④ Search all possible network topology structure and learn relevant parameters based on the training samples. And then we can get the corresponding training accuracy of the different topology structures;

- ⑤ Single out the topology structure of the best training accuracy as the final training results;

- ⑥ In terms of the training results, we can compute the posterior probability of a new (unknown class label) sample X belonging to each class C_i , $P(C_i | X_1, \dots, X_n)$, $1 \leq i \leq n$, by virtue of the formula (9) and (1), where n denotes the number of class. In the end, the sample X is labeled as belonging to the class C^* if $C^* = \max_i \{P(C_i | X_1, \dots, X_n)\}$;

- ⑦ Statistical analysis of classification results.

3. EXPERIMENTS

3.1 Experimental area

GeoEye-1, launched in September 2008, is the latest in a series of commercial high-resolution Earth observation satellites. With its ground sample distance (GSD) of 0.41m for the panchromatic band, GeoEye-1 offers the highest resolution yet available to the spatial information industry. However, for commercial users, image products are down-sampled to 0.5m GSD (Clive S. Fraser, 2009). In the experiment, we choose GeoEye-1 images, which was achieved in 2010 and located in Beijing, China. And the size are 1906 pixel \times 1816 pixel and it covers an area of about 3.5 square kilometres. In order to validate the feasibility of BAN applied in the texture classification of High-resolution satellite images, the above GeoEye-1 images in Beijing are used in the experiments. In fact, there are five classes, such as houses, roads, grass, hills and rivers in the images, which is shown in Figure 2.

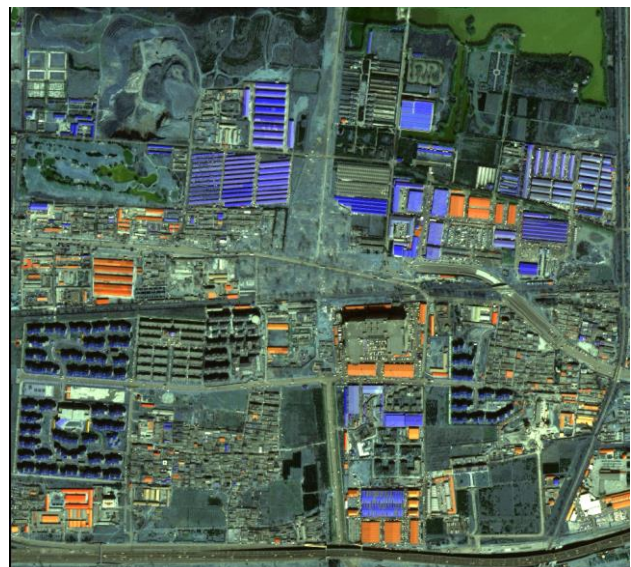


Figure 2. GeoEye-1 High-resolution Satellite Image

3.2 Experimental results

The classification accuracy is calculated based on the confusion matrix, which contains information about the correct classification and misclassification of all classes. To evaluate

the efficiency of BAN, classification results were calculated based on BAN and Naive Bayes Classifiers (NBC) in terms of overall classification accuracy. The experimental results are shown in Table 1.

Table.1 The Comparison results of Two Methods

N	100	150	200	250	300
BAN	82.0	84.4	84.6	86.2	85.8
NBC	80.1	81.3	82.8	82.0	82.0

Table 1 displays comparison of the accuracy among two methods in the condition of different training samples, and N denotes the number of training samples of five classes. The best mean overall classification accuracy is 86.2% (BAN). As expected, BAN gives better classification results than NBC. And Figure 3 is the network topology structure of BAN when the training accuracy is best and the number of training samples is 200. Figure 4 is the classification image based on BAN.

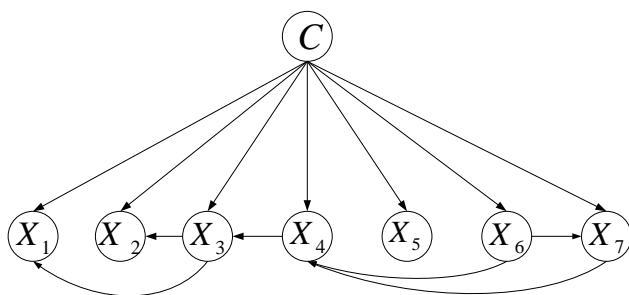


Figure 3. The Topology Structure of BAN

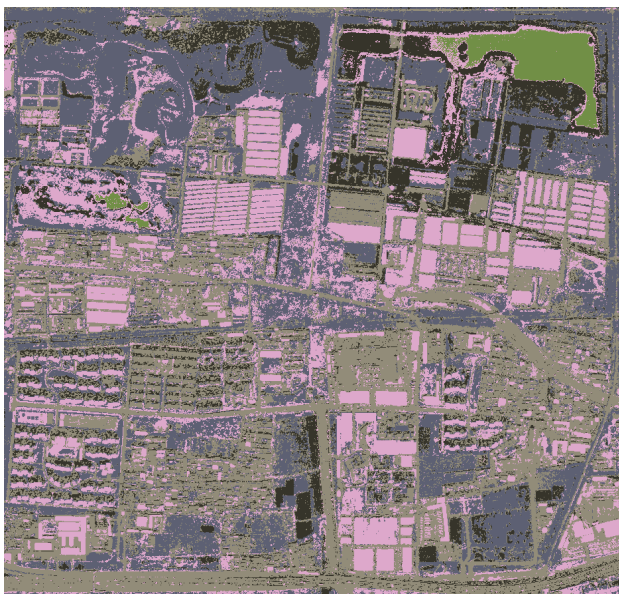


Figure 4. Classification Image Based On the BAN

4. CONCLUSIONS AND FUTURE WORK

Bayesian Network is a directed acyclic graphic model. In this paper, Bayesian Network Augmented Naive Bayes (BAN) is used for texture classification of High-resolution satellite image (GeoEye-1). Experimental results show that BAN outperforms than NBC.

However, search all possible network topology structure needs a great deal of time. In addition, extracting seven kinds of texture features from each classification unit, whose appropriate

window size of high-resolution satellite image we need to study deeply.

ACKNOWLEDGEMENTS

This paper is financially supported by NSFC (No. 40571102). The authors wish to thank the anonymous reviewers for the comments and suggestions on this paper.

REFERENCES

- 1 Friedman, N., Geiger, D., Goldszmidt, M., 1997. Bayesian Network classifiers, *Machine Learning*, vol.29, pp.131~163.
- 2 Langley, P., Iba, W. and Thompson, K., 1992. An analysis of Bayesian classifiers. In *proceeding of AAAI-92*, pp.223-228.
- 3 Yu Xin, Zheng Zhaobao, LI Linyi, Ye ZhiWei., 2005. Texture Classification of High-resolution satellite images based on PCA-NBC, *Proceedings of SPIE - The International Society for Optical Engineering*, v 6044, MIPPR 2005: Image Analysis Techniques.
- 4 D.Heckerman, D.Geiger, D.M.Chickering, 1995. Learning Bayesian networks: the combination of knowledge and statistical data, *Machine Learning* 20, pp.197~244.
- 5 Cheng, J., Greiner, R., 1999. Comparing Bayesian Network Classifiers, In *UAI-99*.
- 6 Cheng, J. & Greiner, R., 2001. Learning Bayesian Belief Network Classifiers: Algorithms and System. *Proc.14th Canadian Conference on Artificial Intelligence*.
- 7 David Heckerman., 1997. Bayesian Networks for Data Mining. *Data Mining and Knowledge Discovery* 1, pp.79~119.
- 8 Jiebo Luo, Andreas E.Savakis, Amit Singhal., 2005. A Bayesian network-based framework for semantic image understanding. *Pattern Recognition* 38, pp.919~934.
- 9 Yu Xin, Zheng Zhaobao, YE Zhiwei, TIAN Liqia., 2007. Texture Classification Based on Tree Augmented Naive Bayes Classifier, *Geomatics And Information Science of Wuhan University*, Vol.32, No.4., p. 287~289.
- 10 Yu Xin, Zheng Zhaobao, Tang Ling, Ye Zhiwei., 2006. High-resolution satellite images Texture Classification Based on Naive Bayesian Network, *Geomatics And Information Science of Wuhan University*, Vol.31, No.2., p. 108~111.
- 11 Cui, Xizhang, Yu, Zhongchou et al., 2001. Generalized adjustment of Survey Measurements. Wuhan: Wuhan University Press.
- 12 Yang S., Chang K., 2002. *IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans*, Vol.32, No.3, pp.419~428.
- 13 Bart B., Geert V., 2004. Bayesian network classifiers for identifying the slope of the customer lifecycle of long-life customers, *European Journal of Operational Research*, vol.156, pp.508~523.
- 14 YU Xin, ZHENG Zhaobao, ZHANG Haitao, YE Zhiwei, 2008. Texture Classification of High-resolution satellite images based on Bayesian network augmented naïve Bayes. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII. Part B7. Beijing, pp.497-501.
- 15 Clive S. Fraser and Mehdi Ravanbakhsh, 2009. Georeferencing Accuracy of GeoEye-1 Imagery, *Photogrammetric engineering & Remote Sensing*, pp.634-638.