International Journal of Engineering Research and General Science Volume 2, Issue 6, October-November, 2014
ISSN 2091-2730

A Survey on Image Classification Algorithm Based on Per-pixel

S.ARUNADEVI[1]Dr. S. DANIEL MADAN RAJA[2]
[1]PG Scholar, Department of IT, Bannari Amman Institute of Technology, Sathyamangalam,
[2]Associate Professor, Department of IT, Bannari Amman Institute of Technology, Sathyamangalam,
arunadevi.se13@bitsathy.ac.in ,daniel@bitsathy.ac.in

Abstract—In this paper, we presents a literature survey on the various approaches used for classifying scenes which is mainly based on object in the given image. In scene classification, classification of images is an intricate process which is the necessity to classify, organize and access them using an easy, faster and efficient way to achieve higher image accuracy within less execution time. The classification of images into semantic categories is an interesting and significant problem. Many different approaches have been proposed relating to object scene classification in the last few years.

Keywords- Image accuracy, Image classification, Supervised classification, Unsupervised classification

I.Introduction

Image classification is an important and challenging task in various application domains, including biomedical imaging, biometry, video surveillance, vehicle navigation, industrial visual inspection, robot navigation, and remote sensing. Classification is an information processing task in which images are categorized into several groups. Categorization of scene allows us to efficiently and rapidly analyze surroundings. A scene is characterized as a place in which we can move. Classifying scenes into semantic categories (such as outdoor, indoor, and sports) is not an easy task. The scene classification problem has two critical components representing scenes and learning models for semantic categories using these representations. When images include occlusion, poor quality, noise or background clutter it is very difficult to recognizing an object in an image and this task becomes even more challenging when an image contain multiple objects.

The main objective of image classification is to identify the features occurring in an image. Supervised classification and unsupervised classification are the two main image classification methods. In supervised classification, trained database is needed and also human annotation is required. In unsupervised classification, human annotation is not required and it is more computers automated. For scene classification many algorithms are referred for classifying the images into semantic categories (e.g. games, sports, street, bedroom, mountain, or coast) [18]. Classification is one of the several primary categories of machine learning problems [6]. The indoor - outdoor scene retrieval problem, how high-level scene properties can be inferred from classification of low-level image features [1]. An automated method has proposed based on the boosting algorithm to estimate image orientations [18]. The classification of indoor and outdoor images based on edge analysis [4]. Analysis of texture requires the identification of proper attributes or features that differentiate the textures of the image [2][6]. For classification of scene images into war scene and nature scene images, the major tasks are identification feature extraction method and suitable classifier. In this paper, presents a literature survey on the various approaches used for classifying images based on Per-pixel Classification

II.Various Classifiacation Methods

There are several ways of grouping the existing scene classification algorithms. Grouping could be based on analyst's contribution in classification methods, or based on parameters on data used, or based on pixel information used, or based on knowledge available from ancillary data, or based on image attributes used. Based or analyst's role, scene classification can be supervised and unsupervised classification. Based on parameters on data used, scenes can be classified as parametric and non-parametric classification. Based on pixel information, scene classification can be per-pixel, sub-pixel, per-field, and contextual classification. Based on availability of knowledge, images can be classified as knowledge-free and knowledge-based classification. In this paper, Scene classification algorithms are described based on Per-pixel Classification techniques in the following subsection

A. Pre-pixel Classification

In per-pixel classification each pixel is assigned to a class by considering the spectral similarities with the different classes [21]. Per-pixel classification can be parametric or non-parametric. In parametric classification it is assumed that the probability distribution of each of the classes is known. Usually, parameters like mean vector and covariance matrix are obtained from the training data. However, the assumption of normal probability distribution of each class is often violated for complex landscapes. Moreover, insufficient training samples may lead to a singular covariance matrix.

The most commonly used parametric classifier is the maximum likelihood classifier (MLC). Unlike parametric classification, non-parametric classification is neither based on any assumption nor uses statistical parameters. This classifier assigns pixels to classes based on pixel's position in discretely positioned feature space [22]. Some of the most commonly used non-parametric classifiers are nearest neighbor (NN), support vector machine (SVM), artificial neural network (ANN) based classifiers, and decision tree-based classifiers.

1) Nearest Neighbor Classification:

Nearest neighbor based algorithms are simple but effective methods used in statistical classification. Categorizing unlabeled samples is based on their distance from the samples in training dataset. Let a set of n labeled training samples be given as $S = \{X_1, X_2, \ldots, X_n\}$, where X_i € R^d. According to the nearest neighbor classification rule, an unlabeled sample t is assigned to the class of X_i € S if X_i happens to be the nearest neighbor of t. Usually Euclidean distance is used as a measure of nearest neighbor. On the other hand, according to kNN classification a set of k nearest neighbors is computed for an unlabeled sample instead of a single nearest neighbor. Then, the test sample is assigned to the class that occurs most frequently among the k-nearest training samples. If the ranges of the data in each dimension vary considerably, this can affect the accuracy of the nearest neighbor based classifications. Thus, both the training and testing data need be normalized [27]

2) Support Vector Machine Classification:

SVM is an efficient supervised binary classification technique. SVM classification methods have often found to provide higher accuracies compared to other methods, such as MLC, ANN-based classifications. SVM classifiers always deliver unique solutions, since the optimality problem is convex. Some of the significant contributions in SVM classification include cluster assumption based active- learning for classifying remote sensing images proposed by Patra et al. [30], fusion of texture and SIFT-based descriptors for remote sensing image classification proposed by Risojević et al. [31], image classification based on linear distance coding proposed by Wang et al. [32]. These algorithms are presented briefly as follows:

Patra et al. [30] develops a reliable active learning based classification for remote sensing images. Collecting labeled samples is time consuming and costly. Also, redundant samples slow down the training process. Thus, training set needs to be kept as small as possible to avoid redundancy, and at the same time, patterns with the largest amount of information need to be included in the training set. The proposed active learning method is implemented in the learning phase of the SVM classifier. The SVM classifier is first trained with a small number of labeled samples. Each unlabeled sample is given an output score based on how likely or unlikely it is a member of a class. These output scores are plotted into a histogram. Thus, the most ambiguous samples generate output scores located in the valley region of the histogram. A threshold is chosen to determine which unlabeled samples should be considered. This technique is not strongly affected by the initial training samples chosen and it is simple in terms of computational complexity. Thus, it has important advantage in remote sensing applications. Risojević et al. [31] proposes a hierarchical fusion of local and global descriptors in order to classify high resolution remote sensing images. They suggest use of a Gabor filter bank at S scales and K orientations. An Enhanced Gabor texture descriptor (EGTD) is developed based on cross correlation between the spatial-frequency sub-bands of Gabor image decomposition.

Wang et al. [32] develops a linear distance coding (LDC) based classification method. Bag of Words (BoW) based classifier uses the three-step method: extraction of local features of an image, generating codebook and then quantize/encode local features accordingly, finally pooling all the codes together to generate a global image representation. However, because of the quantization process, the information loss is inevitable in such a feature extraction-coding-pooling based method. Naïve Bayes Nearest Neighbor (NBNN) method tackles this information loss by avoiding the quantization/coding process. Instead, it uses image-to-class distance, which it calculates based on local features. Since, spatial context of images needs to be explored more effectively for better performance of a classifier, Spatial Pyramid Matching (SPM) is often used as coding-pooling based methods. However, SPM strictly requires that the involved images exhibit similar spatial layout. The proposed method uses the advantages of both BoW and NBNN, and at the same time relieve the strict spatial layout requirement for SPM. In this method each local feature is transformed into a distance vector, whose each element represents certain class-specific semantics. Since image representation produced by LDC is complementary to the one produced by original coding-pooling method, their combination can result in performance improvement of a classifier. Performance is evaluated using both Locally-constrained Linear Coding (LLC) and Localized Soft-Assignment Coding (LSA) as the linear coding method. LLC and LSA are individually used as coding methods, whereas max pooling is always employed. Original coding pooling based image representation, LDC based image representation, and their concatenation are used for evaluation. It is observed that the concatenated representation outperforms the other two in terms of classification accuracy.

3) *Decision Tree-based Classification:*

A supervised classifier which requires less complicated training compared to the ANN is based on a decision tree. A decision tree breaks up a complex decision into multiple simpler decisions so that the final solution resembles the desired solution. Decision tree is a hierarchical structure consisting of nodes and directed edges. Each node is an attribute of an observation that needs to be classified, whereas each edge represents a value the attribute can take. The root node is the attribute, which best divides the training data, whereas each leaf node is assigned a class label. Hunt's algorithm is the most commonly used method for building a decision tree. Hunt's algorithm recursively partitions the training data until all the members of each partition belongs to the same class label, or there are no more attributes remaining for partitioning [37]. Selecting the best split (also known as attribute selection) is a challenging task while building a decision tree and consequently several measures are proposed in literature. The goodness of a split can be measured quantitatively by several metrics, such as information gain, information gain ratio, Gini index etc. While using a large dataset, a decision tree representation can be significantly complex and, hence, classification may suffer from substantial complexity. As a result, a number of pruning methods are employed to reduce the size of the decision tree by removing sections of the tree, which are insignificant in classifying observations. Two of the significant contributions in decision tree based classification are discussed as follows

Pal et al .[38] proposes the use of a univariate decision tree classifier with error based pruning (EBP). They use four different attribute-selection measure metrics to verify that the classification accuracy is not affected by the choice of attribute selection-measure metric. The accuracy of the decision tree classifier is measured while using different pruning methods, such as reduced error pruning (REP), pessimistic error pruning (PEP), error-based pruning (EBP), critical value pruning (CVP), and cost complexity pruning (CCP). It reveals that the EBP outperforms the other pruning methods. They also perform a comparative evaluation between ANN-based classification and the proposed decision tree-based classification. Accuracy and processing time are recorded for both the ANN-based classifier and the decision tree based classifier, using ETM+ and InSAR datasets. It shows that for both the datasets, the decision tree based classifier performs better than the other in terms of both classification accuracy and processing time. Thangaparvathi et al. [39] proposes a modification to the RainForest algorithm, which was developed to address the scalability issue when a large dataset is used. The data structure used in this proposed method IAVC set and IAVC group is the improved version of AVC set (attribute-value class) and AVC group used in the RainForest algorithm.

Several other decision tree-based classifiers have been proposed, which use a variation of the Hunt's algorithm as the decision tree induction method. [40, 41] use the classification based on the ID3 algorithm. [42] Uses the C4.5 decision tree classifier and [43] utilizes the CART based decision tree

4) ***Artificial Neural Network-based Classification:***

ANN is a computational model inspired by the biological neural network. It could be considered as a weighted directed graph in which nodes are neurons and edges with weights are connection among the neurons. Each artificial neuron computes a weighted sum of its input signals and generates an output, based on certain activation functions, such as piecewise linear, sigmoid, Gaussian, etc. It consists of one input layer, one output layer, and depending on the application it may or may not have hidden layers. The number of nodes at the output layer is equal to the number of information classes, whereas the number of nodes at the input is equal to the dimensionality of each pixel. Feed-forward ANN with the back propagation learning algorithm is most commonly used in ANN literature. In the learning phase, the network must learn the connection weights iteratively from a set of training samples. The network gives an output, corresponding to each input. The generated output is compared to the desired output. The error between these two is used to modify the weights of the ANN. The training procedure ends when the error becomes less than a predefined threshold. Then, all the testing data are fed into the classifier to perform the classification.

For very high dimensional data, the learning time of a neural network can be very long, and the resulting ANN can be very complex [33]. Consequently, several ANN-based classification algorithms have been proposed in literature, aiming to minimize the complexity. Both [34, 35] suggest use of adaboost algorithm, i.e., building a strong classifier, using linear combination of several weak classifiers. Both of them use a back propagation learning algorithm. A two-layer-ANN with a single hidden layer is used in [34] as a weak classifier. It uses 50 nodes in the input layer and 25 nodes in the hidden layer. The proposed algorithm works by using weak classifier in a number of iterations (t) and by maintaining a distribution of weights for the training samples. Initially the training samples are assigned equal distribution. However, in subsequent iterations weights of poorly predicted training samples are increased. Finally, the weak classifier finds a weak hypothesis which is suitable for the distribution of the samples at that iteration. A confidence score for the weak hypothesis is also calculated.

AVIRIS data is used in [34] to compare the performance of the proposed algorithm with MLC. It is evaluated that the maximum likelihood-based classifier requires 4,554 parameters for learning, whereas the proposed algorithm requires only 975 parameters for learning but still the proposed method outperforms the maximum likelihood-based classifier. The INRIA human database is used in [35] for evaluation. Three different combinations of weak classifiers are tested by varying the number of nodes in the hidden nodes from 1 to 3. When more hidden nodes are used, the accuracy of the proposed classifier is seen to be better. Comparing the proposed classifier with global linear SVM, global kernel SVM, and cascade linear SVM based classifiers show that the proposed algorithm performs better than the others.

In order to address the complexity of the ANN in case of high dimensional data, feature reduction mechanisms have also been investigated in literature. Majhi et al. [36] proposes a low complexity ANN for recognition of handwritten numerals. The proposed method uses an image database of handwritten numerals. Three hundred ninety six data for each numeral is used for training purposes. At first, binary images of the numerals are converted into gray scale images. Then gradient and curvature values are computed for each image, and subsequently 2,592 dimensional gradient feature vectors and 2,592 dimensional curvature feature vectors are generated. A principal component analysis (PCA) technique is used to compress the data and generate gradient feature vectors and curvature feature vectors of dimensions 66 and 64, respectively. These feature vectors are extended to trigonometric terms and fed to a low complexity single layer classifier. Each numeral with 100 data is used for testing purposes. The classification accuracy obtained using gradient feature vectors and curvature feature vectors are 98% and 94% respectively. Also, it is evaluated that the performance of the proposed algorithm is comparable to the modified quadratic discriminant function (MDQF) based classifier, but it offers low complexity.

III. CONCLUSION

Scene classification plays a key role which is affected by many factors in the field of computer vision. Classification algorithms can be per-pixel, sub-pixel, per-field, contextual, knowledge-based, and high-level. Success of a classification method depends on several factors.Per-pixels classification methods are mostly used in practice. However, they suffer from mixed pixel problem, particularly for medium and coarse spatial resolution data. This paper aims at providing a guide for selecting appropriate classification method based on Per-pixel classification by giving brief knowledge about different classification methods

REFERENCES:

[1] G. Yan, J. F. Mas, B. H. P. Maathuis, Z. Xiangmin, and P. M. Van Dijk, "Comparison of pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China," International Journal of Remote Sensing, vol. 27, pp. 4039-4055, 2006/09/01 2006.

[2] B. R. Kloer, "Hybrid parametric/non parametric image classification," in Technical Papers, ACSM-ASPRS Annual Convention, 1994, pp. 307-316.

[3] M. P. Sampat, A. C. Bovik, J. K. Aggarwal, and K. R. Castleman, "Supervised parametric and non-parametric classification of chromosome images," Pattern Recognition, vol. 38, pp. 1209- 1223, 2005.

[4] S. Patra and L. Bruzzone, "A Fast Cluster-Assumption Based Active-Learning Technique for Classification of Remote Sensing Images," Geoscience and Remote Sensing, IEEE Transactions on, vol.49, pp. 1617-1626, 2011.

[5] V. Risojevic and Z. Babic, "Fusion of Global and Local Descriptors for Remote Sensing Image Classification," Geoscience and Remote Sensing Letters, IEEE, vol. 10, pp. 836-840, 2013.

[6] W. Zilei, F. Jiashi, Y. Shuicheng, and X. Hongsheng, "Linear Distance Coding for Image Classification," Image Processing, IEEE Transactions on, vol. 22, pp. 537-548, 2013.

[7] R. Ablin and C. H. Sulochana, "A Survey of Hyperspectral Image Classification in Remote Sensing."

[8] Q. Sami ul Haq, L. Tao, and S. Yang, "Neural network based adaboosting approach for hyperspectral data classification," in Computer Science and Network Technology (ICCSNT), 2011 International Conference on, 2011, pp. 241-245.

[9] Y. Ren and B. Wang, "Fast human detection via a cascade of neural network classifiers," in Wireless, Mobile and Multimedia Networks (ICWMNN 2010), IET 3rd International Conference on, 2010, pp. 323-326.

[10] B. Majhi, J. Satpathy, and M. Rout, "Efficient recognition of Odiya numerals using low complexity neural classifier," in Energy, Automation, and Signal (ICEAS), 2011 International Conference on, 2011, pp. 1-4.

[11] D. Pop, C. Jichici, and V. Negru, "A combinative method for decision tree construction," in Symbolic and Numeric Algorithms for Scientific Computing, 2005. SYNASC 2005. Seventh International Symposium on, 2005, p. 5 pp.

[12] M. Pal and P. M. Mather, "A comparison of decision tree and backpropagation neural network classifiers for land use classification," in Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International, 2002, pp. 503-505.

[13] B. Thangaparvathi, D. Anandhavalli, and S. Mercy Shalinie, "A high speed decision tree classifier algorithm for huge dataset," in Recent Trends in Information Technology (ICRTIT), 2011 International Conference on, 2011, pp. 695-700.

[14] T. Srinivasan, M. Sathish, V. G. Krishna, and V. Krishnamoorthy, "TC-ID3: A TESTCODE Based ID3 Classifier for Protein Coding Region Identification," in Computational Intelligence for Modelling, Control and Automation, 2006 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on, 2006, pp. 110-110.

[15] F. Yang, H. Jin, and H. Qi, "Study on theapplication of data mining for customer groups based on the modified ID3 algorithm in the ecommerce," in Computer Science and Information Processing (CSIP), 2012 International Conference on, 2012, pp. 615-619.

[16] A. Taherkhani,"Recognizing sorting algorithms with the C4.5 decision tree classifier," in Program Comprehension (ICPC), 2010 IEEE 18th International Conference on, 2010, pp. 72-75.

[17] Z. Zili, Q. Qiming, G. Junping, D. Yuzhi, Y.Yunjun, W. Zhaoqiang, et al., "CART-Based Rare Habitat Information Extraction For Landsat ETM+ Image," in Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International, 2008, pp. III-1071-III-1074.

[18] Andrew Payne and Sameer Singh.2005. Indoor vs outdoor scene classification in digital photographs, Pattern Recognition, pp. 1533-1545.

[19] Arivazhagan S, Ganesan L. 2003. Texture Segmentation Using Wavelet Transform. Pattern Recognition Letters, pp. 3197– 3203.

[20] Bosch A, Zisserman A. 2008.Scene classification using a hybrid generative/discriminative approach, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 30, no. 4, pp. 712–727.

[21] Chang T, Kuo C,1993.Texture Analysis and classification with tree structured wavelet transform, IEEE Transactions on Image Processing, Vol. 2, No.4, P. 429-441.

[22] Lei Zhang, Mingjing Li, Hong-Jiang Zhang, 2002. Boosting Image Orientation Detection with Indoor vs. Outdoor Classification, Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision, December 03-04, pp. 95