International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-2 Number-3 Issue-5 September-2012

A Review of Content Based Image Classification using Machine Learning Approach

Sandeep Kumar[1], Zeeshan Khan[2], Anurag jain[3]
Department of Computer Science & Engineering, Radharaman Institute of Technology and Science Bhopal[1,2,3]

Abstract

Image classification is vital field of research in computer vision. Increasing rate of multimedia data, remote sensing and web photo gallery need a category of different image for the proper retrieval of user. Various researchers apply different approach for image classification such as segmentation, clustering and some machine learning approach for the classification of image. Content of image such as color, texture and shape and size plays an important role in semantic image classification. but the proper selection of feature are challenging task of classification, so various authors apply some machine learning approach for image classification such as decision tree, RBF network, Markova model and support vector machine. In this paper we review of machine learning approach for image classification.

Keywords

Image classification, Support Vector Machine, Feature Extraction, Machine Learning

1. Introduction

Content-based image classification is aimed at efficient classification of relevant images from large image databases based on automatically derived imagery features. These imagery features are typically extracted from shape, texture, color properties of query image and images in the database. Potential application includes digital libraries, commerce, Web searching, biomedicine, surveillance, geographic information systems and sensor systems, education, commerce, crime prevention, etc. Image class refers to the labeling of images into one of a number of predefined categories [1,2]. Although this is usually not a very difficult task for humans, but it has proved an extremely difficult problem for machines. Major resources of difficulty include variable and sometimes uncontrolled imaging conditions, hard-to-describe and complex objects in an image object; objects occlude other objects, and gap between arrays of numbers representing physical images and conceptual information perceived by humans [16, 17]. Designing automatic image class algorithms has been an important research field in recent decades. Potential applications include Web searching, digital libraries, geographic information systems, biomedicine, surveillance, commerce, sensor systems, and education. In terms of classification, image class can be applied as a preprocessing stage: grouping images in the database into semantically meaningful categories. Within the areas of pattern recognition, image processing and computer vision; there has been abundance of prior work on recognizing, detecting, and classifying a relatively small set of objects or concepts in specie domains of application. In the role of classification features extraction play a important role. Lower features content classified image such as color texture and dimensions. For the proper assignment of class of features used machine learning approach [18]. In this approach two event are occurred one is iteration process and another one is statistical approach. In recent trend support vector machine is important tools for image classification. For the diversity of support vector machine also suffered some problem such as outlier problem and core point problem. Now in this problem reduced by other research using some optimizations technique and improve the rate of classification of data.

State-of-the-art image machine learning methods require an intensive learning/training stage. In contrast, non-parametric Nearest-Neighbor (NN) based image classifiers require no training time and have other favorable properties. However, the huge performance gap between these two families of approaches rendered NN based image classifiers and decision tree or support vector machine [20].

The rest of paper is organized as follows. In Section II discuss image features extraction technique. The Section III discusses machine learning technique. Section IV presents the comparative result of methods followed by a conclusion in Section V.

2. Feature Extraction of Image

Feature extraction process plays an important role in content based image classification [5]. In digital image basically three types of features are color, texture and dimensions. Feature extraction can be defined as the act of mapping the image from image space to the feature space. Now days, finding good features that effectively represent an image are still a difficult task[9]. In this section a wide variety of features are used for image classification from the database. Image content can differentiate between visual and semantic content. Features basically represent the visual content. Visual content can be further divided into general or domain specific. For example the features that can use for searching would be representing the general visual content like color, shape and texture. On the other hand, the features that are used for searching human faces are domain-specific and may include domain knowledge. If we talk about the semantic content of an image is not easy to extract. Histogram is important method for color features extraction here we used histogram extraction [11]. Color Histogram Considering a three-dimensional color-space (x, y, z) quantized on each component to a finite set of colors which correspond to the number of bins Nx, Ny, Nz, color of the image I is the joint probability of the intensities of the three color channels. Let i 2 [1,Nx], j 2 [1,Ny] and k 2 [1,Nz]. Then, h(i, j, k) = Card{p 2 I | color(p) = (i, j, k)}. The color histogram H of image I is then defined as the vector H(I) = (..., h(i, j, k), ...). Good performances in texture discrimination and segmentation, the justification for Gabor filters are also supported through psychophysical experiment. Texture analyzers implemented using 2-D Gabor functions produce a strong correlation with actual human segmentation [3]. Gabor functions are Gaussian modulated by complex sinusoids. In the two dimensions they take the form:

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right) \text{...(1)}$$

A dictionary of filters can obtained by appropriate dilatations and rotations of g(x,y) generating function:

$g_{mn}(x,y) = a^{-m} g(x',y')$ *where m=0,1,…..,S-1*

$x' = a^{-m}(x\cos\theta + y\sin\theta)$, $y' = (-x\sin\theta + y\cos\theta)$ ….(2)

Where μ = n¼/K, K the number of orientations, S the number scales in the multi resolution, and a = (Uh/Ul)−1/S−1 with Ul and Uh the lower and upper center frequencies of interest. Compact representation needs to be derived for learning and classification purposes. Given an image I(x, y), its Gabor wavelet transform is then defined as:

$$W_{mn}(x,y) = \int I(x,y) g_{mn} * (x - x_1, y - y_1) dx_1 dy_1 \text{ ..(3)}$$

Where* represents the complex conjugate. Mean μ_{mn} and the standard deviation ¾mn of the magnitude of the transform coefficients are used to represent the image.

$$\mu_{mn} = \iint |W_{mn}(x,y)| dxdy$$

and $\sigma_{mn} = \sqrt{\iint (|W_{mn}(x,y)| - \mu_{mn})^2 dxdy}$ ..(4)

Then a feature vector is constructed using μ_{mn} and σ_{mn} as feature components:

$f = [\mu_{00}\ \sigma_{00}\ \mu_{01}\ \sigma_{01}\ \ldots\ \mu_{mn}\sigma_{mn}]$ ... (5)

As result, we obtain a numerical vector of 30 dimensions for 6 orientations and 5 scales changes. Also note the texture feature is computed only for rectangular grid as it is difficult to compute the texture vector for one arbitrary region.

3. Machine Learning Technique

Machine learning technique improves the performance of image classification. In this section discuss three machine learning technique.

A. Decision Tree

Decision tree ensemble methods are very popular in machine learning. These methods rely in improving an existing learning algorithm by combining the predictions of several models. They are more effective when used with decision trees that otherwise are often not competitive in terms of accuracy with other learning algorithms. We think ensemble methods based on decision trees are a good starting point for designing a generic system for image classification. They are not made any a priori assumption about the application problem; they have been successfully applied to many complex problems in various application domains (see e.g. [2] for some recent applications) and, they compare very favorably with other state-of-the-art algorithms. In this scenario, we first propose to apply a particular ensemble method for decision trees to image classification problems. The method consists in building many extremely randomized trees. It was first proposed in [1]. Extremely randomized trees have the same structure as classical decision trees [4] but the induction algorithm is different. In the classical induction algorithm, tree is grown in a top-down fashion by using the learning examples and searching at each node for the test that maximizes score measure information that evaluates the ability of a test to separate instances of the current learning subset. on the converse, in the extremely randomized induction algorithm, a tree is grown by selecting at each node the test attributes fully at random and its threshold is chosen randomly around the mean of its current values. In the context of image classification, this yields the simple recursive function. Where LS is initialized with all the learning example. Testing at the internal nodes are of the form $[a_{k,l}<a_{t,h}]$ that compare the value of the pixel at the position(K,l) to a threshold a_{th}. Several extra-trees are then built from the same learning sample (in practice as many as possible) and to made prediction for an image, we propagate successfully the entire image into all the trees and we assign to the image the majority class among the classes given by the trees. Below here some steps of algorithm.

Build _extra _tree (LS):
1. If LS contains all images of the same class then return a leaf with this class associated to it;
2. Otherwise:
I.Set $[\alpha_{k,l}<\alpha_{th}]$=Choose a random_ split(LS);
II Split LS into LS_{left} and LS_{right} according to the test $[\alpha_{k,l}< \alpha_{th}]$ and build the subtreesT_{left} build extra tree(LS_{left}) and T_{right}= build _extra_ tree(LS_{right}) from these subsets;
3. Create a node with the test $[\alpha_{k,l}<\alpha_{th}]$attach T_{left} and T_{right} as successors of this node and return the resulting tree.
Choose _a _random _split(LS):
1. Select a pixel location (K,l) at random;
2. Select a threshold a_{th} random according to a distribution $N(\mu_{k,l},R_{k.l}]$, where $\mu_{k,l}$ and $R_{k.l}$ are respectively the mean and standard deviation of the pixel values $\alpha_{k,l}$ in LS ;
3. If the score of this test is greater than a given threshold S_{th} return the test $[\alpha_{k,l}<\alpha_{th}]$;
4.Otherwise, return to step 1 and selecta different location. If all locations already have been considered thenreturn the best test so far.

B. Support Vector Machine

SVMs are learning systems that use a hypothesis space of linear functions in a hyperspace [14], trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory. The aim of Classification via SVM is to find a computationally efficient way of learning good separating hyper planes in a hyperspace, where 'good' hyper planes mean ones optimizing the generalizing bounds and by computationally efficient' we mean algorithms[16,17] able to deal with sample sizes of very high order. The basic problem that a SVM learns and solves is a two-category classification problem. Follow the method of Bennett's discussion (2000), suppose we have a set of *l* observations. Every observation can be represented by a pair $\{x_i, y_i\}$ where $x_i\epsilon\ R^N$ and $y_i\epsilon\ \{-1, 1\}$. That is, each observation contains an N-dimensional vector *x* and a class assignment *y*. Our aim is to find the optimal separating hyperplane; that is, the flat (N-1)-dimensional surface that best separates the data. For time being we assumed that a separating hyperplane exists, and is defined by normal vector *w*. On the either side of this plane we construct a pair of parallel planes such that:

$w\cdot x_i \geq b + 1 \qquad \text{for } y_i = 1$
$w\cdot x_i \leq b - 1 \qquad \text{for } y_i = -1$

where,*b* indicates the offset of the plane from the origin. This Often, a non-linear solution plane is required to separate data. To repeat the steps and maximize the separation between two non-linear functions can be computationally expensive[12,8]. Instead, the kernel trick is used: input data are mapped into a higher dimensional feature space via a specified kernel function. These data are linearly separable in the higher dimensional space. A method of accommodating errors and outliers in the input data was developed , and can be implemented simply by allowing an error of up to ξ in each dimension (resulting in a 'fuzzy margin') and adding a cost function C(*i*) to the optimization equation (Burges). We then want to minimize:

$½\|w\|^2 + C\cdot(\Sigma\ \xi_i)$

Subject to the constraint:

$y_i\ (w\cdot x_i - b) + \xi_i \geq 1$

This is substantially harder to solve than the separable case. In the Chang and Lin's LIBSVM manual, the constraints, minimization conditions and resulting decision functions are defined for each type of classification, along with algorithms to solving the required quadratic programming problems.

C. KNN-Classification

KNN-image classifiers provide good image classification when the query image is similar to one of the labelled images in its class. Actually, NN-image classifiers have proved to be highly competitive in restricted image classification domains, where the number of labelled database images is very high relative to the class complexity. With a theoretical point of view, NN classification tends to Bayes optimal classifier as the sample size tends to infinity [9,10]. However, NN-image classifiers cannot generalize much beyond the labelled image set. In many practical cases, total number of samples (the number of training/labelled images) is very small relative to the class complexity.

When there are only few labelled images for classes with large variability in object shape and appearance, bad classification is obtained. When images are represented by "bag-of-features" histograms so "Image-to-image" distance becomes the 'distance' between two descriptor distributions of the two images (which can be measured via histogram intersection, Chi-square, or KL-divergence)[21]."Image-to-Image" KL-distance (divergence) involves measuring the average log-likelihood of each descriptor d ∈ I1 given the descriptor distribution in I2 [6,7]. Consequently, NN-Image classifiers employ the descriptor distribution of each individual image I ∈ C separately. Instead, we used the descriptor distribution of the entire class C (using all images I ∈ C), we would get better generalization capabilities than by employing individual "Image-to-Image" measurements. Like a direct "Image-to-Class" distance can be obtained by computing the KL-distance between the descriptor distributions of Q and C. even though the "Query to-Image" KL-distance is large for all the 'labeled' images in the Ballet class, the "Query-to-Class" KL-distance may be still small, enabling correct classification.

4. Implementation Details

In this section we implement three methods for image classification using machine learning technique .KNN, Decision Tree and support Vector machine. We evaluated performance of our algorithm using a general-purpose image database containing 500 JPEG images with size of 256*256 or 256*384 pixels from COREL photo gallery. These images are divided into4 categories, and there are 100 images in each semantic category. We test the performance of; the retrieval performance is measured by precision and recall, they are defined below.

$$\text{Precision}=\frac{\text{number of relevant images retrieved}}{\text{number of images retrieved}} \quad \text{...... (1)}$$

$$\text{Recall}=\frac{\text{number of relevant images retrieved}}{\text{number of relevant images in database}} \quad \text{....(2)}$$

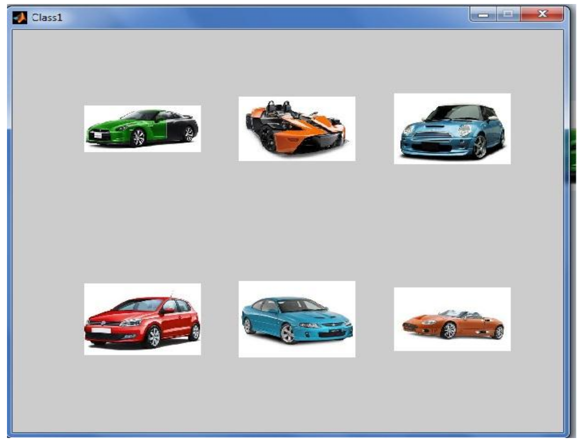

Fig. 1: shows the classified image by machine learning technique

Fig. 2: shows the classified image of dataset for SVMOn the basis of existing methods of image classification Decision Tree, Support Vector Machine and KNN we categorized the Result

Table 1: Result comparison

	Method	**Accuracy (%age)**	**Precision**	**Recall (%age)**
Data-Set 1	DECSION TREE	92.40	87.24	84.43
	SVM	97.14	96.11	94.10
	KNN	91.01	86.12	82.45
Data-Set 2	DECSION TREE	89.90	84.32	82.23
	SVM	95.23	92.14	91.21
	KNN	84.56	83.67	81.34

Data-Set 3	DECSION TREE	91.34	86.14	85.11
	SVM	95.12	93.21	91.13
	KNN	90.10	83.67	82.11
Data-Set 4	DECSION TREE	92.22	88.21	87.66
	SVM	97.13	94.52	93.67
	KNN	89.02	80.67	79.90

Afterwards we have seen in Table1 each Data-Set of image we getting better result in SVM method in terms of accuracy, precision and recall.

Fig. 3: shows that comparative result analysis of DT, SVM and KNN

Here in Fig.3 we have shown that the result in the graph for each Data-Set which is clearly visible in terms of accuracy, precision, and recall SVM produce better result than other.

5. Conclusion

In this paper we perform a comparative result analysis of SVM, DT and KNN classification for image classification. We saw that SVM classifier perform better result in comparisons of other technique. But SVM also suffered from features outlier problem and core problem. In future we remove such problem using optimization of data processing technique using genetic algorithm.

References

[1] M. Kim, M. Madden, and T. A. Warner, "Forest type mapping using object-specific texture measures from multispectral IKONOS imagery: Segmentation quality and image classification issues," Photogramm. Eng. Remote Sens., vol. 75, no. 7, pp. 819–829, Jul. 2009.

[2] Y. O. Ouma, J. Tetuko, and R. Tateishi, "Analysis of co-occurrence and discrete wavelet transform textures for differentiation of forest and nonforest vegetation in very-high-resolution optical-sensor imagery," Int. J. Remote Sens., vol. 29, no. 12, pp. 3417–3456, Jun. 2008.

[3] M. Simard, S. S. Saatchi, and G. De Grandi, "The use of decision tree and multiscale texture for classification of JERS-1 SAR data over tropical forest," IEEE Trans. Geosci. Remote Sens., vol. 38, no. 5, pp. 2310–2321, Sep. 2000.

[4] M. Pesaresiand and J. A. Benediktsonn, "A new approach for the morphological segmentation of high-resolution satellite imagery," IEEE Trans. Geosci. Remote Sens., vol. 39, no. 2, pp. 309–320, Feb. 2001.

[5] J. A. Benedktsonn, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," IEEE Trans. Geosci. Remote Sens., vol. 43, no. 3, pp. 480–491, Mar. 2005.

[6] M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," IEEE Geosci. Remote Sens. Lett., vol. 8, o. 3, pp. 542–546, May 2011.

[7] M. Dalla Mura, J. AtliBenediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 10, pp. 3747–3762, Oct. 2010.

[8] D. Tao, X. Tang, X. Li, and X.Wu, "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 7, pp. 1088–1099, Jul. 2007.

[9] G. Guo, A. K. Jain, W. Ma, and H. Zhang, "Learning similarity measure for natural image retrieval with relevance feedback," IEEE Trans. Neural Netw., vol. 13, no. 4, pp. 811–820, Jul. 2002.

[10] V. Vapnik, The Nature of Statistical Learning Theory. NewYork: Springer-Verlag, 1995.

[11] S. Tong and E. Chang, "Support vector machine active learning for image retrieval," in Proc. ACM Multimedia, 2001, pp. 107–118.

[12] J. Li, N. Allinson, D. Tao, and X. Li, "Multitraining support vector machine for image retrieval," IEEE Trans. Image Process., vol. 15, no. 11, pp. 3597–3601, Nov. 2006.

[13] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, and S. Lin, "Graph embedding and extensions: A general framework for dimensionality reduction," IEEE Trans. Pattern Anal. Mach.Intell., vol. 29, no. 1, pp. 40–51, Jan. 2007.

[14] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," Science, vol. 290, no. 22, pp. 2323–2326, Dec. 2000.

[15] J. Tenenbaum, V. Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, vol. 290, no. 22, pp. 2319–2323, Dec. 2000.
[16] D. Tao, X. Tang, and X. Li, "Which components are important for interactive image searching?," IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 1, pp. 3–11, Jan. 2008.
[17] M. Belkin and P. Niyogi, "LaplacianEigenmaps and spectral techniques for embedding and clustering," in Proc. Adv. NIPS, 2001, vol. 14, pp. 585–591.
[18] J. Chen, S. Shan, G. Zhao, X. Chen, and W. Gao, "MattiPietikainen, WLD: A robust descriptor based onWeber's law," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1705–1720, Sep. 2010.
[19] C. Hoi, C. Chan, K. Huang, M. R. Lyu, and I. King, "Biased support vector machine for relevance feedback in image retrieval," in Proc. IJCNN, Budapest, Hungary, Jul. 25–29, 2004, pp. 3189–3194.
[20] X. He, D. Cai, and J. Han, "Learning a maximum margin subspace for image retrieval," IEEE Trans. Knowl. Data Eng., vol. 20, no. 2, pp. 189–201, Feb. 2008.
[21] W. Bian and D. Tao, "Biased discriminant Euclidean embedding for content-based image retrieval," IEEE Trans. Image Process., vol. 19, no. 2, pp. 545–554, Feb. 2010.

I am **Sandeep Kumar**. Student of M.Tech dept. computer science & engg. RITS, Bhopal (India).My interest area is Image processing and neural networking.