IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Texture based Ranking of Categories in a Natural Image

Janhavi H. Borse[1*], Dipti D. Patil[2]

[1]Department of Information Technology, Sandip Institute of Technology & Research Centre, Nashik, India
[2]Department of Information Technology, MKSSS's Cummins College of Engineering for Women, Pune, India

[*]*Corresponding Author: borse.janhavi@gmail.com, Tel.: +91-9225367554*

DOI: https://doi.org/10.26438/ijcse/v7i5.183187 | Available online at: www.ijcseonline.org

Accepted: 14/May/2019, Published: 31/May/2019

Abstract— Natural scene images are captured at a larger distances to include details in scenery. It is much difficult to identify categories because of uncertain shapes & forms present inside these images. Such ambiguous form of nature, which lacks sharp boundaries, makes discrimination among the classes a complex task. This paper attempts to measure this ambiguity. A natural scene image also can belong to multiple categories at a time which makes a task of classification much more difficult and often leads to classification errors. Binary classification fails to capture this ambiguity while doing multi label classification of the image. This problem can be handled by using fuzzy membership function with assumption that class categories in a natural image are non-mutually exclusive. This work provides a ranking based class membership instead of binary classification.

Keywords—Fuzzy Membership Function, Multi-Label Classification, Ranking, Supervised Learning

I. INTRODUCTION

A natural scene image consists of many categories exhibiting different semantic meanings and hence they can be discriminated. It means that all the patterns belonging to a particular category, although they are different, follow similar characteristic behavior. Depending on the above discussion some may conclude that the natural scene images are conveying multiple semantic meanings and hence their classification task is a multi label one. An image can be assigned a single label or multiple labels depending upon its characteristics. Many of the approaches [1], [10], [12], [13] & [14] deal with scene understanding problem as multi label classification problem. All those approaches try to solve the problem based on the assumption that scene categories are mutually exclusive. It is equivalent to the task of assigning a single image to multiple class labels, which seems to be quite illogical and unnatural as the real scenario is exactly opposite.

In reality scene categories are overlapped with each other, which depict ambiguity in the nature [1]. Hence it is more logical to rank the objects in the images according to their degree of membership. The boundaries in natural image are fuzzy than in case of synthetic images. They are uncertain due to presence of non-mutually exclusive classes in a scene image. In the literature it is found that some uncertainties can be represented using probability theory [1]. The ambiguity can be best modeled by using fuzzy membership functions. It specifies the degree of membership of an object to a particular class label. As scene images are captured from large distances to fit the entire scene in the images, most of the details are lost. Hence the color component, shape and size of natural objects cannot be quantified. So what remains is the texture component and hence in case of natural scene images it is better to assume that if an image is represented and described by its texture, it would result in better discriminative analysis of its objects.

In supervised learning, input space is characterized by most discriminating image features. Its corresponding feature vector represents an input image. An output space is the set of labels out of which any subset may be associated with an unseen image depending on its input characteristics. The task is to model a classifier function by analyzing the training samples from input data for which associated set of labels are known. The trained classifier is then used to classify unseen image and predict set of relevant labels for it.

Existing system with relative attributes faced a challenge of lesser accuracy. We aim to improve the accuracy of this system by employing texture features and by re- training the classifier with adding each query image to the dataset after predicting its labels.

Section II includes a related work for different feature extraction methods, Classifiers in the literature for scene classification problem. Section III includes implementation details of the proposed system along with technical terms. Section IV includes experimental setup & results. Section V concludes the paper.

II. RELATED WORK

It is found that statistical methods like co-occurrence matrix has major difficulties involving high time complexity if an image is represented with high number of intensity levels. Model based methods like fractals [2] can efficiently describe roughness in natural scene images. But as natural surface is not deterministic but would always have some statistical variation, it makes the computation of fractal dimension much more difficult. All the sinusoidal transforms & Laws' [3] mask provide comparable results when comparisons are done using misclassification probabilities. It seems that if textures are used for describing images, Gabor filter would be a better choice. But if implementation is concerned Laws' masks are better as its misclassification probability is negligible although more than that of Gabor filters. The work [1], [4] used Gabor wavelet filters for feature extraction, analysis and com- pared their results with other multi-resolution texture features for image retrieval problem. It is found that the results with Gabor are more robust than the others mentioned earlier. Methods [1], [5] evaluated texture feature extraction operators using number of filters. The filters used are derived from discrete transform, Gabor filters & Laws' masks. It is evident that greater the separation, better is the classification. A survey [6], [1] gives detailed analysis of various texture based segmentation methods such as statistical methods, geometrical methods, signal processing methods like spatial domain filtering, Fourier transform filters, Fourier domain filters, Gabor & wavelet filters etc. Spatial filtering methods such as LOG can work on many textures and can discriminate both natural & synthetic textures by controlling some parameters of estimation. Fourier domain also gives similar results. Gabor & wavelet models integrate frequency analysis into spatial domain thereby localizing the global frequency analysis. Also it is found that integrating a region based method with boundary based method obtain more robust and clean segmentation. Approaches as described in [6], [7], [8], [15] give details about texture segmentation and classification based on Gabor filters, filter selection, Feature selection and computation of efficient texture features for image classification purpose. Approach [9] tried to focus on the tasks such as feature extraction and representations so that the application areas like object recognition, image classification and content based image retrieval (CBIR) systems would be benefited with fewer efforts. The aim was to bridge the gap between human perception of scene images and their computer vision counterparts.

In the analysis [5], label correlations are used to categorize the multi label methods but still there are no specific benchmarks for exploiting label correlations especially with domains that have large output space. The work [10], [12] compared state-of-the-art classifiers viz. Hierarchy of Multi label classifiers, Random Forest of Predictive Clustering, Classifier Chains (CC) & Binary Relevance (BR). Authors in [11] studied the multi label problem with the assumption that the multi label learning algorithms are categorized depending on the label correlations they use either first order, second order or higher order. It is found that multi label learning models the complex semantics in output space and assumes relevance ordering of each class label such that a binary decision in classification is converted into an ordered membership. Random Forest of Predictive Decision Trees abbreviated as RF-PDT is found to be the better algorithm. Method [13] used BR-KNN as their base classifier & used iterative approach. An approach [14] converted the image classification problem into the optimization problem with objective function using sparseness in label indicator so that images can be classified with respect to relevant or irrelevant labels. Authors in [13] tried to solve the problem of multi-label classification for input set with iterative learning. Authors in [16] tried to tackle the problem of representing uncertainties in nature. They tried to help in selecting the most appropriate fuzzy membership function for scene images. It is useful for representing uncertainties in the nature. An approach [16] considered scene images as containing non-mutually exclusive data. This method integrates fuzzy reasoning with qualitative reasoning and then maps semantics of an image onto the output space of predefined classes. Ranking [1], [16] was according to membership degree of the classes (confidence value) with respect to an image. The classification accuracy was more than 70%. But the overall time required is more as compared to other traditional approaches as it needs to calculate membership value of each feature with respect to each and every class. Although it has measured the ambiguity in natural scene images very precisely, the time complexity needs to be reworked on. This is because it is directly proportional to the number of classes and features used. So if more and more data has to be classified it would require more time which would be hazardous to any real time applications. Work in [17] tried to classify flower images to different floral classes using color information. Approach [18] classified small sized general images into respective classes using convolution neural network and analysed the effect of increasing number of network layers on performance of the network.

Related work [1], [16] on scene understanding has been analyzed and found that an existing system using fuzzy rank classifier face the challenges of time complexity being directly proportional to the number of class labels and number of features used for classification task. The classification accuracy of existing system is also prominently lesser than that using conventional multi- label classifiers like SVM. The existing system does not focus on the features used for classification. It has used relative attributes. If texture features are used instead of relative attributes, accuracy of the existing system can be improved. It is also observed that, the features extracted using Gabor filtering technique, are more efficient in a sense that such features would help improve classification accuracy.

III. METHODOLOGY

The system is trained using training database and fuzzy membership matrix M is generated as the output. In testing phase, an unseen query image will be taken as an input from user and the system produces ranked classes as an output. For both the phases, feature extraction is of equal importance.

3.1 Feature Extraction:

The proposed system uses a bank of 2-dimentional Gabor filters with varying orientations centered at number of the most dominant frequencies. Features are generated from the outputs obtained by filtering input images using this bank. Input to this stage will be an image $I_k(x , y)$ from the set of N number of images denoted & defined by $I = \{I_1, I_2,....,I_N\}$. The response of the filter is denoted & defined by,

$$h(x,y)=g'(x,y)\exp(j(\omega_x x+\omega_y y)) \quad (1)$$

where,

$$g'(x,y) = \frac{1}{\lambda\sigma^2} g\left(\frac{x'}{\lambda\sigma}, \frac{y'}{\sigma}\right) \text{ and}$$

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

$$x' = x\cos\theta + y\sin\theta,$$

$$y' = -x\sin\theta + y\cos\theta$$

The response h(x,y) is a product of a Gaussian low pass filter and a complex exponential. σ is a spatial scaling parameter which is used to control the filter response width. λ is aspect ratio and θ is the orientation angle. On filtering an input image I(x,y) by response h(x,y), the resulting image matrix is represented by 2-dimensional convolution and features are generated from the resultant output samples. The absolute values of intensities of the images so obtained are used to compute the features like mean, standard deviation, average output energy, average contrast between each pixel pairs, and entropy. In all 50 features are obtained, using the 10 selected filters and 5 features for each filter response.

3.2 Classifier Training:

The classifier will be learned with the help of a fuzzy membership function. The membership value of each feature for each class will be approximated by 4-tuple fuzzy number, m = {α, a, b, β}. The fuzzy representation [1] describes the gradual change in the membership degree and hence it can be used to better quantify a quality of a natural scene.

For every feature and for every class x_{jk}, a histogram of frequency of occurrence is calculated to obtain m_{jk}. The mutually exclusive region can be located using a threshold μ which is calculated as,

$$\mu = \frac{\sum_{i=1}^{B} n_i}{B} \quad (2)$$

,where B are predefined number of bins of histogram and n_i are the number of training images that satisfy the range of values for respective bins. An unknown image, I_Q is given as an input to the trained classifier and its features will be extracted. For feature j of class k the membership value μ is approximated by,

$$\mu_{jk}(x_j) = \begin{cases} 0, & x_j < a - \alpha \\ \alpha^{-1}(x_j - a + \alpha), & a - \alpha \leqslant x_j < a \\ 1, & a \leqslant x_j \leqslant b \\ \beta^{-1}(b + \beta - x_j), & b < x_j \leqslant b + \beta \\ 0, & x_j > b + \beta \end{cases} \quad (3)$$

Rank of a class is calculated by multiplying all values across all rows of matrix μ_{JK}.

3.3 Algorithm

We assume that the system is trained using aforementioned classifier training section and we have matrix M_{JK} .

Input: Unseen Image $\mathbf{I_k(x,y)}$

Output: Relevant class categories **C** of image I_k ranked in the order of their confidence levels.

Steps:

1. For each feature of each class repeat the steps 2 to 4
2. Build histogram for each feature from x_{jk}, which represents number of occurrences of training images in respective bins. Number of bins B is empirically set to 60.
3. Compute mean value for each histogram built in step 2
4. Use a threshold equal to the mean value obtained from step 3 for defining dominant & overlapped regions for corresponding feature and class.
5. Return μ_{JK} containing the values α, β as cut off for overlapped region and a, b as lower & upper bounds of the dominant region for corresponding feature and class.

6. Compute Rank using,
   $R_k = \frac{P_k}{\sum P_k}$ and $P_k = \prod_{j=1}^{J} \mu_{jk}(x_j)$ is a Product P_k of membership values is calculated across all the features for each class k.

IV. RESULTS AND DISCUSSION

The Outdoor Scene Recognition (OSR) data set is used to test the performance of the system. It is a collection of 256 X 256 color images of natural scenes. Dataset includes 8 outdoor scene classes. It contains 2688 labeled images in all. Figure 1 shows sample images from each class.

The system is tested to identify the number of correctly classified images and number of incorrectly classified images over the entire database. The overall system accuracy is thus calculated using,

$$\text{Accuracy} = \frac{\text{Number of Images correctly classified}}{\text{Total Images in Database}}$$

The following system parameters are set empirically. Number of bins, B = 60, Orientation angle $\theta = \{0^0, 60^0, 90^0, 120^0, 150^0\}$

Figure 1. Sample images in OSR dataset.

Through this experiment number of correctly classified and incorrectly classified images over the entire database are observed. In correct classification the images are divided into two types: I. Images with class membership greater than or equal to 0.5 II. Images with class membership less than 0.5 but still it is the highest among all. In the incorrectly classified prediction the membership of the true class is not the highest. Table 1 shows the Overall Accuracy for the OSR dataset, where 2566 images from 8 categories were applied as query. Average System Accuracy is 97.3 %. The lowest accuracy of 95 % is observed for coast category.

Table 1. Overall Accuracy for OSR dataset

Classes	**Correctly Classified**		***Incorrectly Classified C***	***Accuracy*** $\frac{A+B}{A+B+C} \times$ **100**
	A (membership >= 0.5)	***B (membership < 0.5)***		
Coast	***250***	***20***	***12***	***95***
Forest	290	25	10	96.9
High-way	239	13	0	1
Inside-city	281	15	12	96.1
Mountain	318	44	8	97.2
Open Country	327	48	6	98.4
Street	242	44	6	97.9
Tall Building	328	17	11	96.9
Total	2275	226	65	97.3

The class membership for unseen query images is shown in figure 2. Almost all unknown images are correctly classified with ground truth class membership more than 0.95. Unlike other classes, image with ground truth of Inside City is incorrectly classified as Tall Building with membership of 0.835. It still has Inside City class membership of 0.164, which shows that the classifier is confused between the two classes viz. Inside City & Tall Building. It shows realistic prediction of the classifier using fuzzy membership function.

V. Conclusion and Future Scope

Existing system assumed that any kind of the features could be used for classification. But we assumed that for successful classification efficient features are needed. For this reason texture features are used which are extracted using Gabor filter. Because of using texture features instead of relative attributes the performance of the existing system has been increased. Almost 97% of the images from OSR dataset are correctly classified as per the ground truth. The number of images, which are incorrectly classified, is also very small. It means that the erroneous predictions have been reduced.

The future work may include testing the system on different databases with large number of images with many categories. Further it can be studied to account for the effects of different class membership functions on the system.

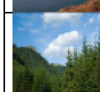
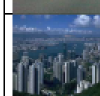
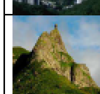
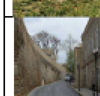
Image	Classes	Class Membership Grade							
		Coast	Forest	Highway	Inside City	Mountain	Open Country	Street	Tall Building
	Coast	0.981	0	0.015	0	0.001	0	0	0.001
	Forest	0	1	0	0	0	0	0	0
	Highway	0	0	0.972	0.027	0	0	0	0
	Inside City	0	0	0	0.164	0	0	0	0.835
	Mountain	0	0	0	0	1	0	0	0
	Open Country	0.264	0.057	0	0	0.067	0.61	0	0
	Street	0	0	0	0	0	0	1	0
	Tall Building	0	0	0	0	0	0	0	0.999

Figure 2. Class memberships for unseen query images.

Acknowledgment

It was a support Colleagues and family, who knowingly or unknowingly helped us in the successful planning, designing

& developing phases of this work. I am glad to express my gratitude to all those who rendered their valuable time and helped for the successful completion of the project.

REFERENCES

[1] Janhavi Borse, N. M. Shahane, "*Multi-Label Classification of A Scene Image using Fuzzy Logic*", *IJCA Proceedings ETC 2016*, ISBN : 973-93-80975-01-2, vol. 01, no. 2, pp. 4-10, March 18, 2017.

[2] Alex P. Pentland, "*Fractal-Based Description Of Natural Scenes*", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. Pami-6, No. 6, pp. 661-674, November 1984.

[3] G. Lemaitre and M. Rodojevi, "*Texture segmentation: Co-occurrence matrix and Laws' texture masks methods*", pp. 1-34.

[4] B.S. Manjunathi and W.Y. Ma, "*Texture Features for Browsing and Retrieval of Image Data*", IEEE transactions on pattern analysis and machine intelligence, vol. 18, no. 8, pp. 837-842, august 1996.

[5] S. E. Grigorescu, N. Petkov, and P. Kruizinga , "*A comparative study of filter based texture operators using Mahalanobis Distance*", 0-7695-0750-6/00, *IEEE*, pp. 1-4, 2000.

[6] M. Tuceryan and Jain , "*Texture Analysis*", ResearchGate Article, pp. 1-42, September 2000.

[7] M. Lindenbaum and R. Sandler, "*Gabor Filter Analysis for Texture Segmentation*", pp. 1-58, May 2005.

[8] J. Ilonen, J.-K. Kamarainen, H. Kalviainen, "*Efficient computation of Gabor features*", pp. 1-29, 2005.

[9] Y. Alqasrawi, D. Neagu and P. I. Cowling, "*Fusing integrated visual vocabularies-based bag of visual words and weighted colour moments on spatial pyramid layout for natural scene image classification*", research gate article in signal image and video processing, pp. 1-25, July 2011.

[10] G. Madjarov, DragiKocev, DejanGjorgjevikj and S. Dzeroski, "*An extensive experimental comparison of methods for multi-label learning*" , Elsevier publication on Pattern Recognition, pp. 3084-3104, 2012.

[11] M. Pagola, C. Lopez-Molina, "*Interval Type-2 Fuzzy Sets Constructed From Several Membership Functions: Application to the Fuzzy Thresholding Algorithm*", IEEE transactions on fuzzy systems, vol. 21, no. 2, pp. 230-244, April 2013.

[12] Min-Ling Zhang and Zhi-Hua Zhou, "*A Review on Multi-Label Learning Algorithms*", IEEE transactions on knowledge and data engineering, vol. 26, no. 8, pp. 1819-1837, august 2014.

[13] J. Wu1, V. Sheng2, J. Zhang3, Peng peng Zhao1, Z. Cui, "*Multi-label Active Learning for Image Classification*", *ICIP*, pp. 1-5, 2014.

[14] L. Jing and M. K. Ng , "*Sparse Label-Indicator Optimization Methods for Image Classification*", IEEE transactions on image processing, vol. 23, no. 3, pp. 1002-1014, march 2014.

[15] M. Celuszak and D. Jabry, "*ESGI100: Gabor Filter Selection and Computational Processing for Emotion Recognition*", pp. 1-23, 18 may 2014.

[16] C. H. Lim, A. Risnumawan, and C. S. Chan, "*A Scene Image is Nonmutually Exclusive-A Fuzzy Qualitative Scene Understanding*", IEEE transactions on fuzzy systems, vol. 22, no. 6, pp. 1541- 1556, December 2014.

[17] Riddhi H.Shaparia, Narendra M.Patel, Zankhana H. Shah, *"Flower Classification using Different Color Channel*", International Journal of Scientific Research in Computer Science and Engineering, Vol.7, Issue.2, pp.1-6, 2019.

[18] N.S. Lele , "*Image Classification Using Convolutional Neural Network*", International Journal of Scientific Research in Computer Science and Engineering, Vol.6, Issue.3, pp.22-26, 2018.

Authors Profile

Mrs. Janhavi H. Borse (Formerly Janhavi D. Shirke) is currently working as Assistant Professor in Department of Information Technology, Sandip Institute of Technology & Research Centre, Nashik. She pursed UG & PG in Computer Engineering from K.K. Wagh College of Engineering & Research Centre, Nashik. Her research interests are Machine Learning, Pattern Recognition, Image Processing. She is a professional member of ACM since 2014 and Life member of ISTE.

Dr. Dipti D. Patil is currently working as Associate Professor in Department of Information Technology at MKSSS's Cummins College of Engineering for Women, Pune. She earned her doctorate in the year 2014 from Sant Gadgebaba Amravati University in Computer Science and Engineering. She pursued her UG and PG Computer Engineering from Thadomal Shahani Engineering College, Bandra of Mumbai University in 2002 and 2008 respectively. She is a Life member of CSI and ISTE. She has authored books in areas of Data Structures and Mobile healthcare. Dr. Dipti has published many research articles, which are published in various national and international Journals and conferences. She is involved in developing healthcare system prototypes and for which she has filed various national and international patents. Her areas of special interest include mobile healthcare, analytics, data science, artificial intelligence and Internet of Things.

