Multi-Label Classification of A Scene Image using Fuzzy Logic

Janhavi Shirke
Student
Department of Computer Engineering
K.K.W.I.E.E.R., Nashik

N. M. Shahane
Associate Professor
Department of Computer Engineering
K.K.W.I.E.E.R., Nashik

ABSTRACT
A natural scene image contains object categories which form ambiguous boundaries. Measuring this ambiguitywhile classifying an image, is a challenging task. A scene image belongs to multiple categories at a time which makes a task of classification multi label one.Binary classification fails to capture this ambiguity while classifying the scene image into one of mutually exclusive classes. This problem can be handled by applying fuzzy logic with non-mutually exclusive class categories. This project work provides a ranking based on class membership instead of binary classification.

General Terms
Image Processing, Pattern Analysis and Machine Vision.

Keywords
Fuzzy logic, Binary classification, Ranking.

1. INTRODUCTION
A scene imageconveys multiple semantic meanings through different class categories and hence classification of a scene image task is a multi label task. An image can be assigned a single label or multiple labels depending upon its characteristics and member classes. Many approaches [9,11,12,13] deal with scene classification problem as multi label classification problem. These approaches try to solve the problem based on the assumption that scene categories are mutually exclusive. It means that the classes form crisp boundaries without overlapping. It is equivalent to the task of assigning multiple class labels to an image which makes the task of classification binary one.

In reality scene categories are overlapped with each other which preserve ambiguity in the nature. Hence it is more logical to rank the categories in the images according to their degree of membership rather than assigning multiple class labels to an image.

For example, for an image with two classes, mountain & ocean, it is hardly possible to locate the boundaries of each class representing the two objects (see Figure 1). This is because the boundaries are no longer crisp as in case of synthetic images.Rather they are uncertain due to the presence of overlappedcategories in a scene image.

In the literature it is found that some uncertainties can be represented using probability theory[19]. The ambiguity in a scene image can be best modeled by using fuzzy membership functions. It specifies the degree of membership of a class to an image.

(a) (b) (c)

Figure 1. Example of single class images with classes (a)coast (b)mountain, overlapped scene classes (c)coast& mountain.

Asa scene image is captured from largerdistance to fit the entire scene in the image, most of the color details are lost. Further a scene can take any natural shape which gets reflected into its image and is uncertain. Hence the color component, shape and size of natural objects cannot be quantified. Hence it is better to assume that if a scene image is described by its texture component, it would be beneficial for classification purposes.

The paper is organized as follows. Section 2 enlists the methods in the literature which tried to solve the scene understanding problem & includes a survey of different feature extraction methods. Section 3 details about the proposed system &its working. Section 4 details the experiments performed and the results. Section 5 concludes the paper.

2. RELATED WORK
2.1 Feature Extraction Methods
Model based methods like fractals [1] 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.

It is found that all the sinusoidal transforms & Laws' [2] 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.

B.S. Manjunathi and W.Y. Ma [3] used Gabor wavelet filters for feature extraction, analysis and found that the results with Gabor filter are more robust than other multi-resolution texture features for image retrieval problem.

S. E. Grigorescu, N. Petkov [4] 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 separability, better is the classification possible.

Mihran Tuceryan and Jain [5] 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. It is found that integrating a region based method with boundary based method obtain more robust and clean segmentation.

Approaches [6], [7] details about texture segmentation and classification based on Gabor filters while approach [14] focuses on filter selection parameters and approach [17] details about computation of efficient texture features for image classification purpose.

Yousef and Peter [8] integrated visual vocabularies (histograms) of all the classes instead of using traditional universal vocabularies like BOWs (bag of words).

2.2 SceneClassification Methods

Gjorgji Madjarov and Dragi Kocev [9] found that Hierarchy of Multi label classifiers abbreviated as HOMER was best when evaluated on recall and Random Forest of Predictive Clustering Trees abbreviated as RF-PCT was best when evaluated on precision (exact predictions). Further Classifier Chains (CC) & Binary Relevance (BR) are ranked next according to their performances evaluated on selected measures.

Miguel and Carlos [10] tried to tackle the problem of representing uncertainties in nature. They tried to help in selecting the most appropriate fuzzy membership function [18] for scene images which could be useful for representing uncertainties in the nature.

Min-Ling and Zhi-Hua [11] found that multi label learning models the complex semantics in label 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.

Jian and Victor [12] used BR-KNN as their base classifier. They trained classifier using labelled set and randomly choose some images from unlabeled set as test data, performed labelling on them and then put those newly tested images into trained set. With this newly formed training set they re-trained the classifier.

Liping and Michael [13] converted the image classification problem into the optimization problem.

There is only one approach by Lim and Chan [15], to the best of our knowledge, in which scene images are considered as containing non-mutually exclusive data. They have integrated fuzzy reasoning with qualitative reasoning and then mapped semantics of an image onto the output space of predefined classes. They ranked the classes according to their membership degree (confidence value) with respect to an image. They proposed Fuzzy Qualitative Rank Classifier (FQRC).Their results are found to be more towards reality with classification accuracy more than 70%. But the overall time required is more as compared to other traditional approaches.

Richard Cabral and Torre [16] tried to solve the problem of multi-label classification for input set with some of the labels missing for some images and converted the multi-label classification problem into the rank minimization problem. The results found were almost comparable to the most powerful linear SVM classifier.

2.3 Motivation

Related work on scene understanding is 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 [15] 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 [15], accuracy of the existing system can be improved. It is also observed that, the features extracted using Gabor filtering technique, are more efficient meaning that these features would help improve classification accuracy [18].

3. TECHNICAL DETAILS

The proposed system is classifying a query image into its relevant label set which are ranked according to their membership degrees using fuzzy membership function and Gabor filter. The system works in three phases: Feature Extraction, Training & Testing.

3.1 Mathematical Model

Let,

$I = \{I_1, I_2, ..., I_N\}$ denotes a set of N scene images.

$X = \{X_1, X, ..., X_N\}$denotes a set of feature vectors,where each X_k denotes J-dimensional feature vector for k^{th} image.

$X_k = \{x_{k1}, x_{k2}, ..., x_{kJ}\}$

$Y = \{1, 2, ..., K\}$ denotes the set of K class labels.

$F = \{F_1, F_2, F_3, F_4, F_5\}$ is a set of functions.

F_1 is used for feature extraction,

F_2 is used for building class histogram for each feature,

F_3 is used for approximating fuzzy membership,

F_4 is used for computing membership degree of relevant classes for the query image,

F_5 is used for calculating rank of each relevant class for the query image.

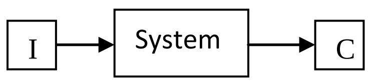
I_Q is the unseen query image given by user.

Output is given as a set of relevant class labels with membership grades,

$C = \{C_1, C_2, ..., C_m\}$,where m is the no. of predicted labels and m < K.

The system is defined as,

$S = \{I, X, Y, F, I_Q, C\}$



3.2 Feature Extraction

The feature extraction is done using Gabor Filter with input filter parameters set according to [14].

Orientation angle $\theta = \{0^0, 60^0, 90^0, 120^0, 150^0\}$

Filter size $\sigma = \{3, 13\}$, wavelength $\lambda = \{4, 8\}$

The absolute values of intensities of the images obtained as a response of each applied filter are used to compute the features as mean, standard deviation, average output energy of filter response, average contrast between each pixel pairs, and entropy. In all 50 features are obtained, using the 10 selected filters for each angle & scale and 5 features for each filter response.

The feature extraction algorithm can be summarized as follows:

1. Group all images according to their known category
2. Input Image I(x,y)
3. Set Gabor filter parameters orientation angle θ, Filter size σ and wavelengthλ and generate one filter
4. Filter the input image with new generated filter from step 3
5. Calculate the features mean, variance, energy, contrast and entropy from the absolute values of filter responses obtained from step 4
6. Repeat steps 3 to 5 for each new filter
7. Add all filter responses

3.3 Training Phase

The classifier is trained using Gabor features computed in feature extraction phase(see figure 2). The membership matrix containing each value equal to 4-tuple fuzzy number, m = {a, b, α, β} for each feature of each class is approximated by a histogram H_{jk}= {h_1, h_2,....,h_B }.$H_{jk}$is a histogram containing no. of occurrences of training images in respective bins for j^{th}feature and k^{th}class, where B is predefined no. of bins and empirically it is set to 60. The range (a-b) represents dominant class region for a particular feature while values of α, β are calculated to represent overlapped class region.

The fuzzy representation describes the gradual change in the membership degree and hence it is used to better quantify a quality of a natural scene.

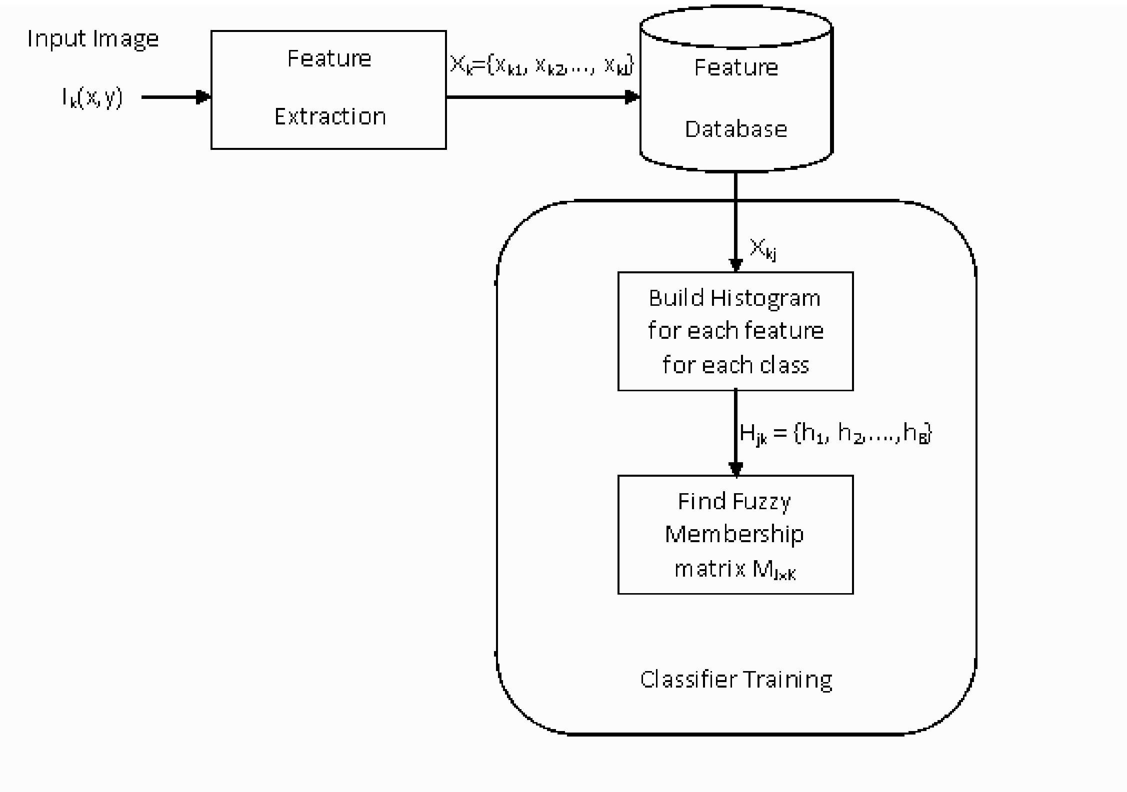


Figure 2. Training Phase

3.4 Testing Phase

In testing phase, a query image I_Q is given as an input to the trained classifier and membership degree of each class is calculated using fuzzy membership function,

$$\mu_{jk}(x_j) = \begin{cases} 0,\ b+\beta < x_j < a-\alpha \\ (x_j - a + \alpha)\,\alpha^{-1},\ a-\alpha \le x_j < a \\ (b+\beta - x_j)\beta^{-1},\ b < x_j \le b+\beta \\ 1,\ a \le x_j \le b \end{cases} \tag{1}$$

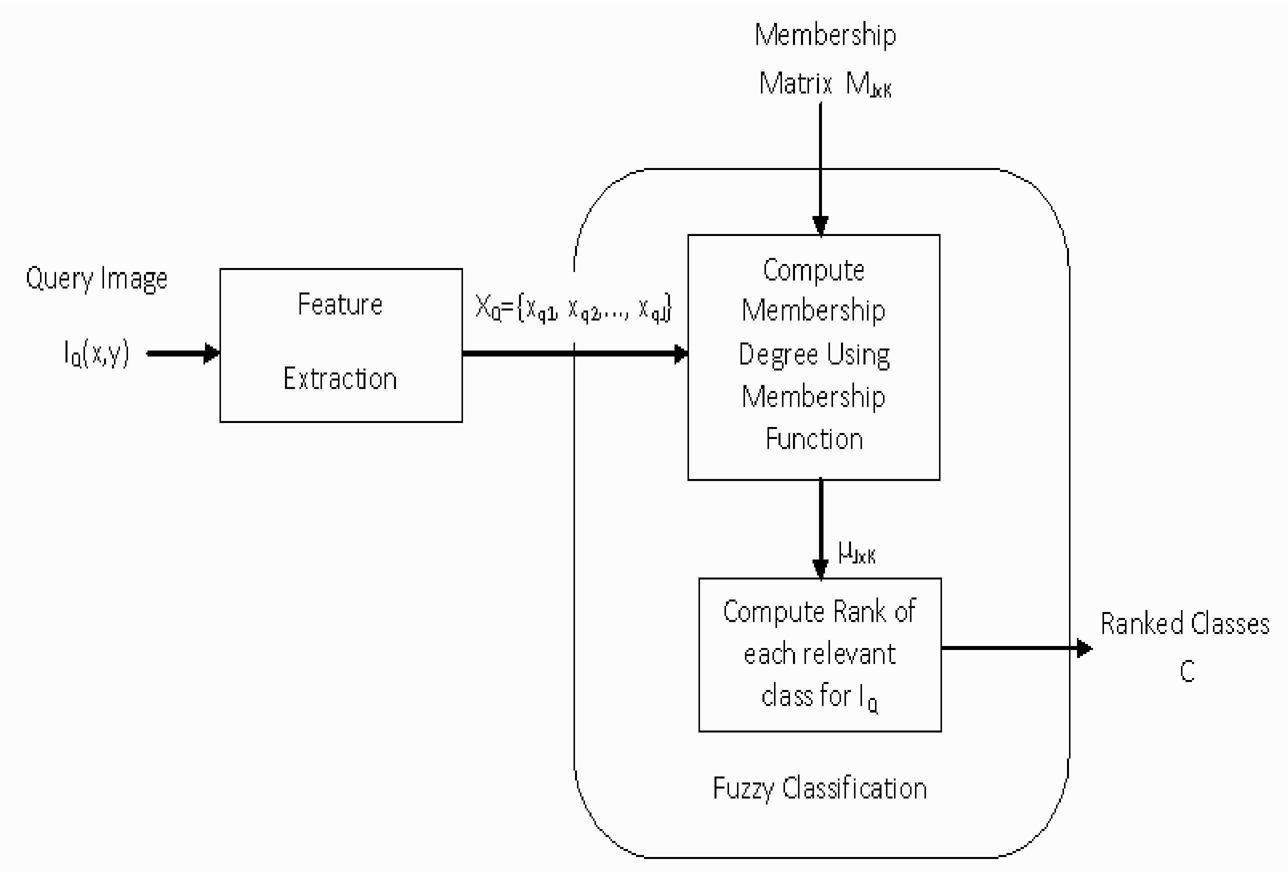


Figure 3. Testing Phase

For a single class rank is calculated by multiplying all values across all rows of matrix μ_{jk}, and then normalizing that value. The process is summarized in figure 3.

4. EXPERIMENTAL SETUP

The Outdoor Scene Recognition (OSR) dataset is used for testing the performance of the proposed system.The OSR dataset 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. Table 1 shows no. of images per class of OSR dataset.

Table 1. No. of images per class of OSR dataset

Coast	**Highway**	**Street**	**Tall Building**
360	160	361	433
Open Country	**Inside City**	**Mountain**	**Forest**
392	280	374	328

4.1 Performance Metric

Accuracy of the proposed system is measured using F-Score,

$$\textbf{F-Score} = \frac{2vp}{v+p} \quad (2)$$

Where ***v*** is recall and ***p*** is precision which are calculated using,

$$v = \frac{true\ positives}{true\ positives + false\ negatives}$$

$$p = \frac{true\ positives}{true\ positives + false\ positives}$$

4.2 Results

In the first experiment, 5 images are chosen randomly from the OSR dataset.Training is done for rest of the dataset. Histogram bins are set empirically to B = 60. Results are obtained for each test image (see Figure 4). Membership Degrees of each class to each test image is enlisted into Table 2.

Table 3 shows the calculated values of precision, recall and F-score for the test images of figure 4. Average F-score can be calculated and is equal to 0.798 which is greater than existing system [15] accuracy without α-cut. After analyzing recall values of five test images (see Table 3), it is clear that none of the ground truth labels is missed.

The time taken by the system to classify a query image varies in the range from 60 to 90 seconds depending on the image. It is to be noted that the results shown are for the first experiment done. Extensive experimentation needs to be performed to test the system accuracy, robustness and reliability.

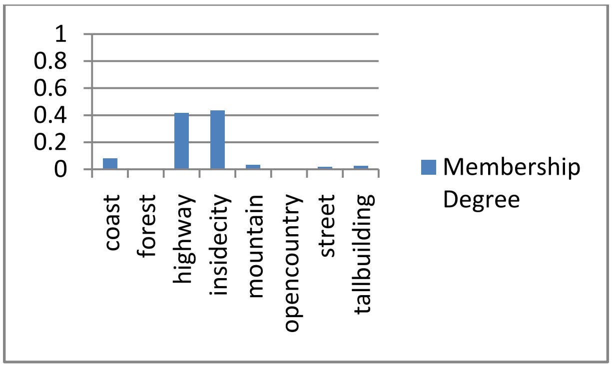
Table 2. Membership Degrees of each class for images of Figure 4

Scene Image	Coast	Forest	Highway	Inside City	Mountain	Open Country	Street	Tall Building
(a)	0.078	0	0.416	0.433	0.0329	0	0.0159	0.023
(b)	0.2628	0	0.0161	0.0084	0.574	0.0056	0.00125	0.1308
(c)	0	0	0	0.9978	0	0	0	0.00153
(d)	0.00854	0	0	0.6559	0.0476	0	0.1524	0.135
(e)	0.00106	0	0	0.992	0.0037	0	0	0.003086

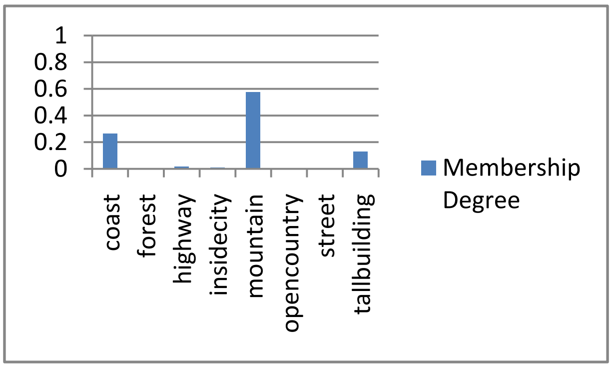
Table 3. Calculated F-score from the experiment

Images	**(a)**	**(b)**	**(c)**	**(d)**	**(e)**
p	0.5	0.71	0.67	0.71	0.75
v	1	1	1	1	1
F-Score	0.67	0.83	0.80	0.83	0.86

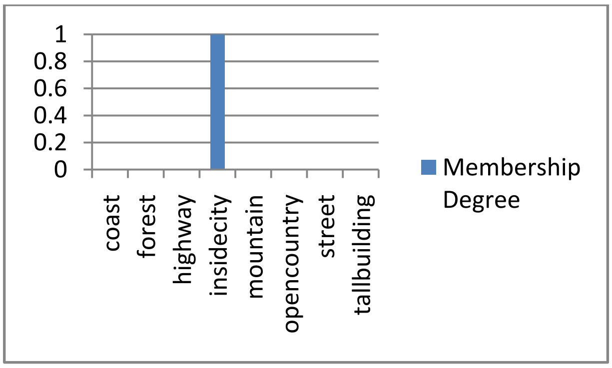
(a) (b) (c) (d) (e)



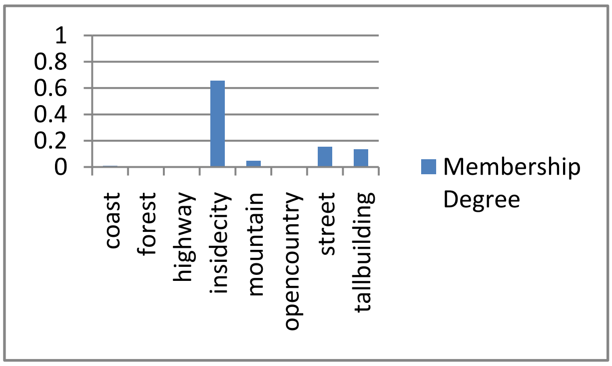
(a)



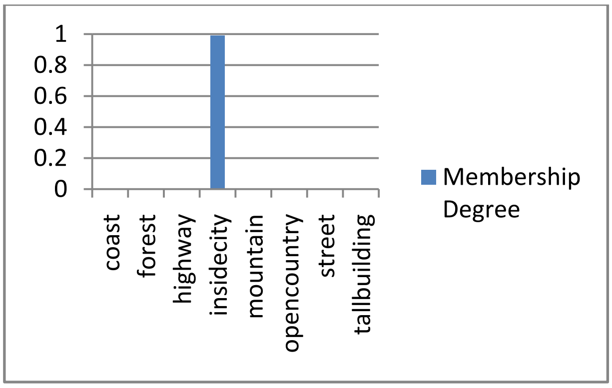
(b)



(c)



(d)



(e)

Figure 4: Test images (a) to (e) and their corresponding class labels with membership degrees

5. CONCLUSIONAND FUTURE SCOPE

The proposed system represents the logical and natural idea of the scene classes being non-mutually exclusive and attempts to measure the ambiguity in natural scene images using fuzzy membership function.It does so by not only predicting the label set for a query image but also showing membership degree of each class to that image.

For the first experiment done the accuracy of the proposed system is greater than the existing system which promotes our assumption of using texture features for image representation.

The fuzzy loss function could be used to redesign the fuzzy logic in future and the experiments can be carried out using different datasets with more number of classes.

6. ACKNOWLEDGMENT

I am glad to express my gratitude to all those who rendered their valuable time and helped for the successful completion of this work. I am thankful to my guide Prof. N. M. Shahane, for his continuous guidance and encouragement. His expert suggestions and scholarly feedback had greatly enhanced the effectiveness of this work.

I am also thankful to Prof. Dr. S. S. Sane, Head, Department of Computer Engineering, K.K.W.I.E.E.R., Nashik, for the interest shown in this work by timely suggestions and helpful guidance. I would also like to express my thanks to Prof. Dr. K. N. Nandurkar, Principal, K.K.W.I.E.E.R., Nashik.

7. REFERENCES

[1] 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.

[2] G. Lemaitre and M. Rodojevi, “Texture segmentation: Cooccurrence matrix and Laws’ texture masks methods”, pp. 1-34.

[3] 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.

[4] 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.

[5] M. Tuceryan and Jain , “Texture Analysis”, *ResearchGate Article*, pp.1-42, September 2000.

[6] M. Lindenbaum and R. Sandler, “Gabor Filter Analysis for Texture Segmentation”, pp. 1-58, May 2005.

[7] J. Ilonen, J.-K. Kamarainen, H. Kalviainen, “Efficient computation of Gabor features”, pp. 1-29, 2005.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] J. Wu1, V. Sheng2, J. Zhang3, Peng peng Zhao1, Z. Cui, “Multilabel Active Learning for Image Classification”, ICIP, pp. 1-5, 2014.

[13] 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.

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

[15] 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.

[16] R. Cabral, Fernando De la Torre, J. Costeira, and A.Bernardino, “Matrix Completion for Weakly-Supervised Multi-Label Image Classification”, *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 1, pp. 121-135, january 2015.

[17] Y. Dong, D.Tao, and X. Li, "Nonnegative Multi-resolutionRepresentation-Based Texture Image Classification", *ACM Transactions on Intelligent Systems and Technology*, Vo. 7, No. 1, Article 4, pp. 1-21, September 2015.

[18] Janhavi Shirke and N.M.Shahane, "Qualitative understanding of a scene image", cPGCON 2016, pp. 1-6.

[19] George J. Klir, Bo Yuan, 'Fuzzy Sets And Fuzzy Logic: Theory and Applications', PHI Learning private limited.

[20] Richard O. Duda, Peter Hart and David Stork, 'Pattern Classification', Wiley India Pvt. Ltd.