Proposed Multi-label Image Classification Method Based on Gabor Filter

Ziad Abdallah, Ali El-Zaart, and Mohamad Oueidat

1 Introduction

Image Multi-label Classification (IMC) is an important topic in data mining that assigns a label or a set of labels to an image. The big demand for image annotation and archiving in the web attracts researchers to develop many algorithms for this application domain [1]. The Multi-Instance Multi-label Learning (MIML) is a framework of machine learning proposed recently for computer vision application [2]. In this framework, an image is described with many regions or instances and can be assigned to multiple labels.

For instance, Fig. 1 shows that the image contains three regions for the label "trees." Each region in the image is a set of instances. These regions can be expressed as different examples called feature vector, and in data mining it is called Multi-Instance Learning. At the same time, the image may be classified simultaneously for more than one label; it is then called Multi-Label Learning. The Multi-Label Learning models the relations between labels and regions (instead of the entire image).

This will decrease noises in this feature space and increase the accuracy of the model [1]. Figure 2 shows that only three regions in the image are assigned to three labels (sky, mountain, and water).

This comparison illustrated that multi-label learning takes into consideration the correlation between labels, while multi-instance learning connects regions to labels. MIML takes both relations simultaneously. MIML has been successfully applied to

Z. Abdallah (✉) · A. El-Zaart
Department of Mathematics and Computer Sciences, Beirut Arab University, Beirut, Lebanon
e-mail: z.abdallah@bau.edu.lb; elzaart@bau.edu.lb

M. Oueidat
Maintenance and Industrial Engineering Department, University Institute of Technology, Saida, Lebanon
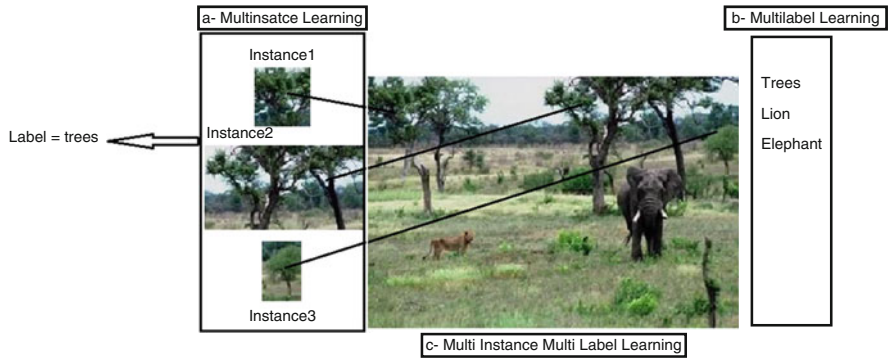
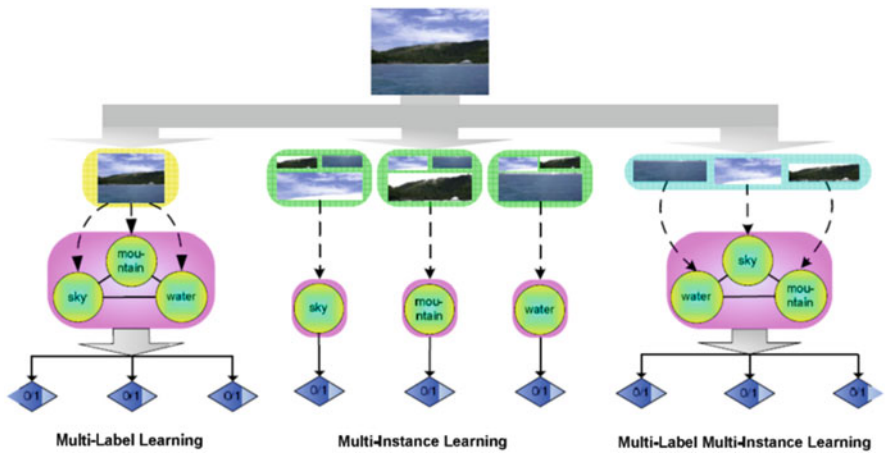**Fig. 1** Multi-Instance Multi-label Learning

Fig. 2 Comparison between the three learnings [3]

image text classification, image annotation, video annotation, ecological protection, and other tasks [1–4].

The first way transforms multi-label to single-label. This transformation is called Multi-Instance Single-label Learning (MISL) and applies multi-instance learner to have Single-Instance Single-label Learning (SISL). The second way transforms multi-instance to single-instance. This transformation is called Single-Instance Multi-label Learning (SIML) and applies multi-label learner to have SISL. Two most important techniques are proposed for these transformations: MIML-Boost and MIML-SVM [4]. We are interested in this chapter in the second transformation as a multi-label problem. The drawbacks of these existing methods do not take into consideration the description of some characteristics from the image and the correlation between labels [5].

This chapter proposes a new framework that improves the MIML. The idea is to extract the feature from the image using Gabor filter bank (GFB). It is a feature extraction algorithm that takes into consideration the local representation, shape, and geometry of an image. Then we apply K-mean to cluster the image into similar groups. The final step consists of applying the Label Priority Power set as multi-label transformation in order to solve the problem of label correlation [6]. Each step of our new contribution is described in Sect. 3.

The remainder of this chapter is organized as follows: Sect. 2 defines the problem formulation. Section 3 defines the problem solution. Section 4 discusses the experimental results. Finally, the conclusions are presented in Sect. 5.

2   Problem Formulation

The MIML is an algorithm that transforms the problem to a single-label classification [4]. In Fig. 3, there are two ways to do this: the first one (T1) transforms MIML to MISL and applies Multi-Instance Learner (L1). The second one (T2) transforms MIML to SIML and applies Multi-Label Learner (L2). A good review can be found in [4]. Two most important techniques are proposed for these transformations:

1. MIML-BOOST: It is an MISL transformation. Each region in the image is transformed into a set of multi-instance called bags. Thus, original data-set can be split into a number of multi-instance datasets with only one label each. The learning task is transformed to traditional single-label learning. Figure 4 shows that the region in the original data-set is assigned to three labels (red green blue) from four (red green blue purple). In this algorithm, the region is split into four multi-instance regions.
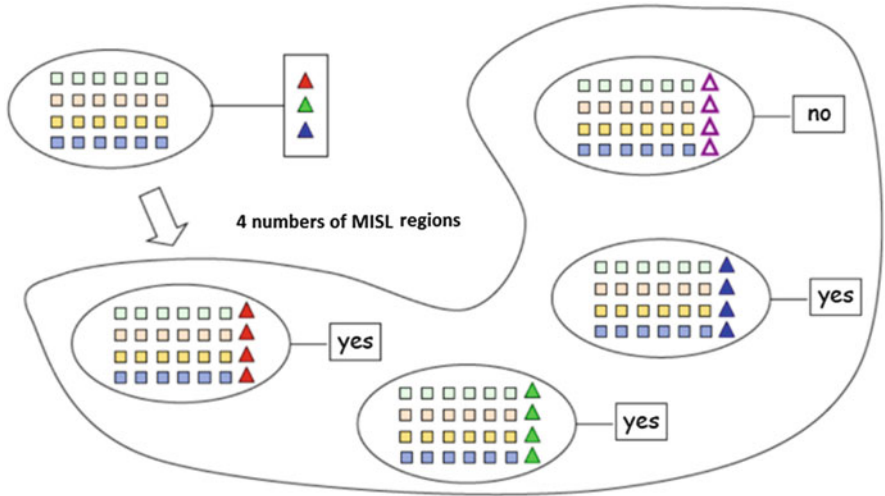

Fig. 3 MIML solutions [4]
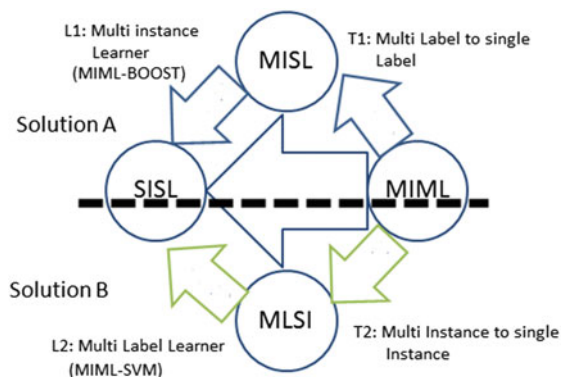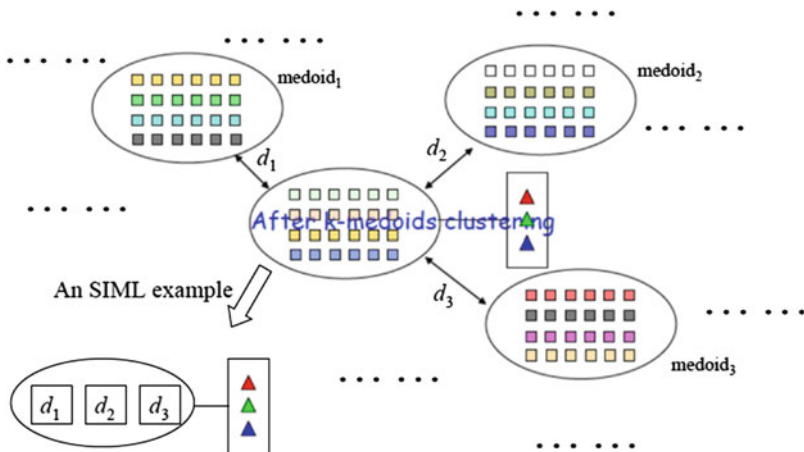
Fig. 4 MIML-boost illustrations [4]

Fig. 5 MIML-SVM illustrations [4]

2. MIML-SVM: It is an SIML transformation. Each object is mapped into a feature vector using Hausdoff distance from a number of medoids generated first. The learning task is transformed to traditional multi-label learning.

Figure 5 shows that the region is mapped into a feature vector with three features computed using Hausdoff distance after using 3-medoids clustering algorithm. Figure 6 illustrates the steps of MIML-SVM.
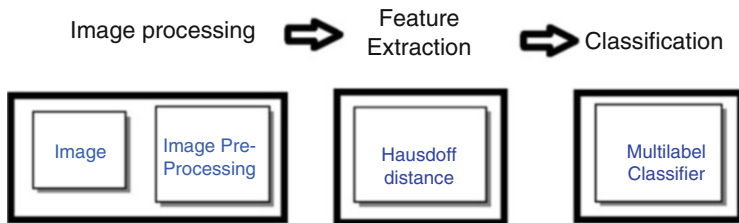
Fig. 6 MIML-SVM steps

The drawbacks of these existing methods are that they do not take into consideration the following:

1. The description of some characteristics from the image: color, shape, regions, textures and motion, and some elementary characteristics. The image could be assigned to sky or lake based on the relative size and placement of the components [5].
2. The correlation between labels: In multi-label learning, labels are correlated. For example, if the mountain label is assigned to the image with rocks and sky and the field label is assigned to image with grass and sky, then an image with grass, rocks, and sky would be assigned with both labels, field and mountain [5].

3 Problem Solution

3.1 Two-Dimensional Gabor Filter

This is a linear filter used in several domains in image processing [7, 8]. Formally, the two-dimensional Gabor filter family is expressed by the following equations:

$$g\left(x, y, \lambda, \theta, \psi, \sigma, \gamma\right) = \exp\left[-\frac{1}{2}\left(\frac{x'^2 + \gamma^2 y'^2}{\sigma^2}\right)\right] \exp\left[i\left(2\pi\frac{x^2}{\lambda^2} + \psi\right)\right] \quad (1)$$

$$x' = x\cos\theta + y\sin\theta \quad (2)$$

$$y' = y\cos\theta - x\sin\theta \quad (3)$$

Four parameters are important to determine the Gabor filter, as shown in Fig. 7, which shows the variation of

1. The wavelength of the sinusoidal factor λ
2. The orientation of the normal to the parallel stripes of a Gabor function θ
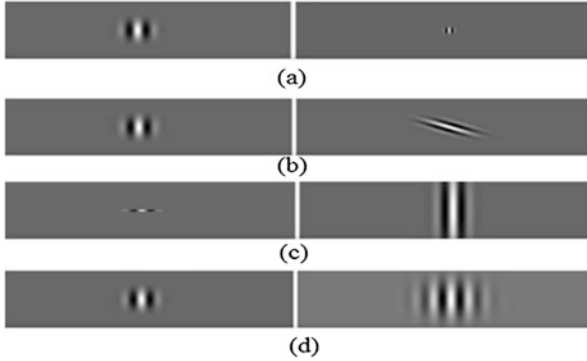
Fig. 7 Variations of the Gabor parameters [9]

Multi-label Dataset D						
Object	**Y**	y_1	y_2	y_3	y_4	**Binary**
E_1	$Y_1 = \{y_1, y_2\}$	1	1	0	0	1100
E_2	$Y_2 = \{y_1, y_2, y_4\}$	1	1	0	1	1101
E_3	$Y_3 = \{y_3, y_4\}$	0	0	1	1	0011
E_4	$Y_4 = \{y_2, y_4\}$	0	1	0	1	0101
Frequency		2	3	1	3	

Multi-label Sorted Dataset D'					
Object	$y'_1 = y_3$	$y'_2 = y_1$	$y'_3 = y_2$	$y'_4 = y_4$	**Binary**
E_1	0	1	1	0	0110
E_2	0	1	1	1	0111
E_3	1	0	0	1	1001
E_4	0	0	1	1	0011
Frequency	1	2	3	3	

Fig. 8 LPP transformation

3. The phase offset ψ
4. γ is the spatial aspect ratio

Moreover, σ is the sigma of the Gaussian envelope and usually it is equals one.

3.2 The Label Priority Power Set (LPP)

The LPP is a transformation method of multi-label learning into SISL [6]. It orders the label by importance. The advantage of this method is that it solves the problem of label correlation [6]. Figure 8 shows the conversion of multi-label data D to a multi-label dataset D' sorted by the frequency of each label.

3.3 Framework MIML-LPPGABOR

This section proposes a new framework that improves the MIML. The idea extracts important features (Mean, Standard Deviation, Skewness, Kurtosis, Entropy, First Quartile, Median, Third Quartile) from image using Gabor filter bank. It is a feature

extraction algorithm that takes into consideration the local representation, shape, and geometry of an image. The first three features are the first central moments. They reflect the center position, the dispersion, and the asymmetry of the probability distribution. The drawback for these three features is that they are sensitive to outliers. Therefore, we add three features that divide the data in the image into four equal groups and they are not sensitive to outliers. Thus, we solve the first and second limitation of MIML. We then apply K-mean to cluster the image into similar groups. The final step consists of applying the Label Priority Power set as multi-label transformation in order to solve the problem of label correlation. The challenge of such learning is that the image contains many concepts existing in several regions at the same time. We faced the following issues:

1. The images do not have the same size [10].
2. Selection of the suitable feature from the image.
3. Different objects in the image could be similar [11].
4. Multiple objects in the same image.

Figure 9 shows block diagram of the new framework compared to MIML (Fig. 6). The following is a detailed representation:

Phase 1: Image Preprocessing
Resizing the image consists of changing the sample rate of the original image, preserving the important content and structure. Formally, let I be an image with m rows and n columns $I_{m \times n}$. The resized image is an image $I'_{m' \times n'}$. The output of this step is a dataset $D(I1_{m' \times n'}, I2_{m' \times n'}, \ldots, Ip_{m' \times n'})$, where p is number of images in the dataset. All images in the dataset D have the same dimension. The advantage of this step is to prepare a dataset for the feature extraction process. The limitation is that only uniform scaling can be applied when resizing the image [10]. Figure 10 shows an example of resizing a sample of images.

Phase 2: Feature Extraction Using Gabor Filter Bank
Gabor filter bank (GFB) is composed of many distinct Gabor filters with different parameters. Two parameters are useful for extracting the suitable features from image [8]: the orientations and the frequencies. They are calculated using the
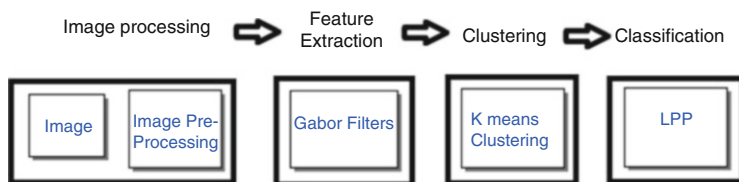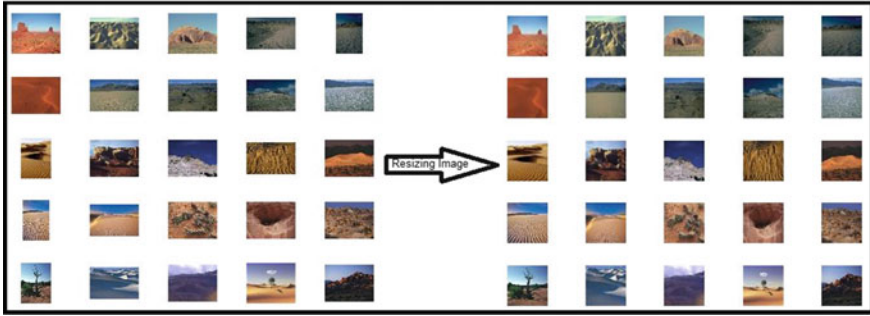

Fig. 9 Block diagram of the proposed method

Fig. 10 Resizing a sample of images

following equations:

$$\theta(i) = \frac{(i-1)\pi}{\Theta}, \quad \text{where } i = 1, 2, \ldots \Theta \text{ (number of orientations)} . \quad (4)$$

$$\omega(i) = \frac{0.25}{\left(\sqrt{2}\right)^{i-1}}, \quad \text{where } i = 1, 2, \ldots S \text{ (number of scales)} \quad (5)$$

Figure 11 shows GFB with five orientations and five frequencies.

The process of extracting feature from image, as shown in Fig. 12, consists of:

1. Reading the original image.
2. Resizing the image and transforming it from RGB to gray space color. The output is an image I with the size ($m = 128$, $n = 128$).
3. Applying each Gabor filter from GFB to I. Formally, this involves convolving each region in the image with the Gabor filter. The output of this step (c) is 25 filtered images with the same size as I.
4. Normalizing each filtered image by zero mean and unit variance. Then, it is subsampled by two factors: $d1$ and $d2 = (4,4)$, as in Fig. 12. That is, meaning that we will select $32 = 128/4$ rows and 32 columns from the image. The output of this step is an image Is with the size (Ms = 32, Ns = 32). Each Is is partitioned into 4×4 Blocks. We extract nine features (Mean, Standard deviation, Skewness, Kurtosis, Entropy, First Quartile, Median, and Third Quartile) from each 2×2 blocks

Phase 3: *K*-Means Clustering

There are two kinds of cluster analysis techniques: *K*-Means and hierarchical Clustering. *K*-Means is better than hierarchical clustering in case of big amount of data [11]. *K*-Means consists of grouping similar images into different k mutually exclusive clusters. The output of this step is K clusters $C1$, $C2$, $\ldots$, Ck. An image
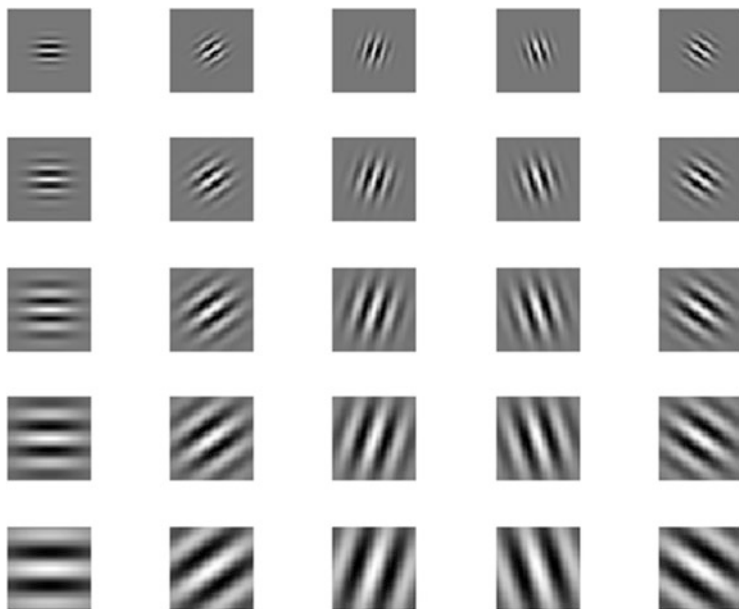
Fig. 11 GFB with five orientations and five frequencies
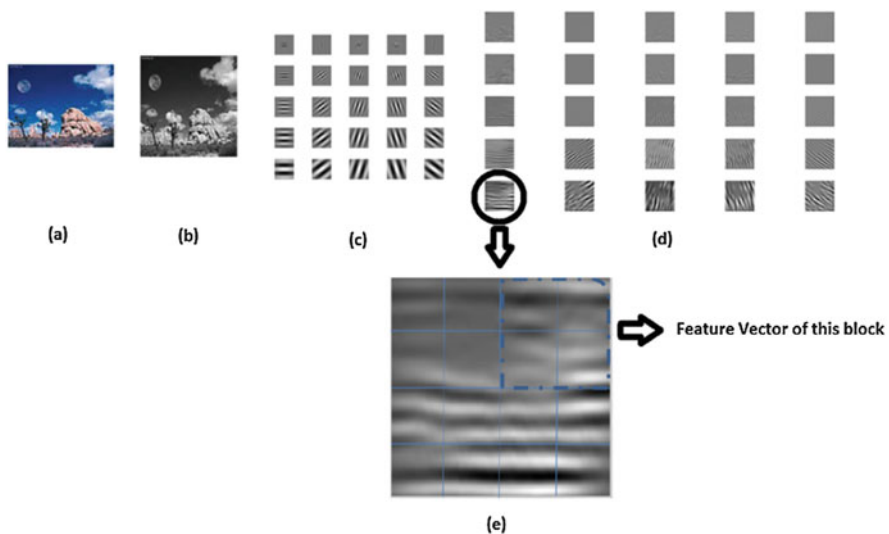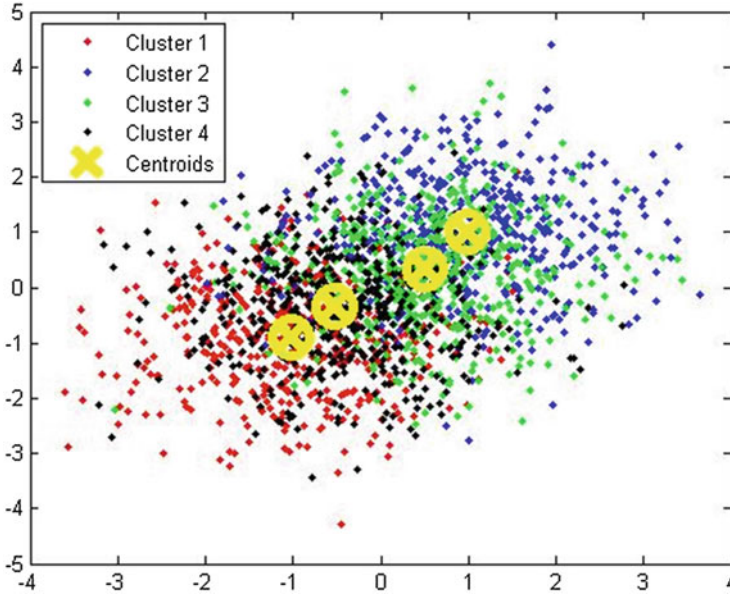

Fig. 12 Feature extraction using GFB

Fig. 13 Clustering method using K-Means

may belong to exactly one of these clusters. The advantage of *K*-Means is that it is better than hierarchical clustering in case of big amount of data when *K* is small. The disadvantage of *K*-Means is the difficulty of predicting the *K* representing the number of clusters. Figure 13 shows the centroids of four clusters generated by the 4-means algorithm.

Phase 4: Converting Multi-label Dataset to Multiclass Dataset using LPP
We will use in this step the transformation problem through breaking down the multi-label dataset into a single-label dataset using LPP transformation [6]. The output of this step is a dataset Ds = {(X1,y1), ..., (XP,yP)}, where *Xi* is the feature extracted from Gabor and *yi* is the decimal conversion of binary multi-label value as shown in Fig. 14. The importance of this step is the reduction of the complexity of learning process.

Tree Decision is a powerful classifier used in this phase because of its ease of use and its independence of the features of the dataset and their distribution. The output of this step is *k* trees, where *k* is the number of clusters. The advantage of this step is to applying single-label classification in a multi-label problem. Figure 15 shows the tree decisions constructed in the training phase for four cluster (the content of each tree is not important in the figure).

Multi-label Sorted Dataset D''

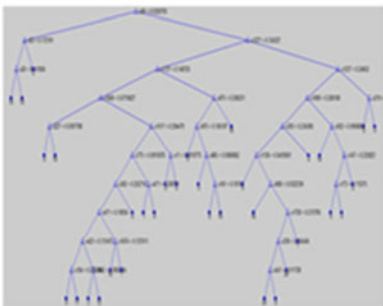| Object	Binary	yDec
E_1	0110	6
E_2	0111	7
E_3	1001	9
E_4	0011	3

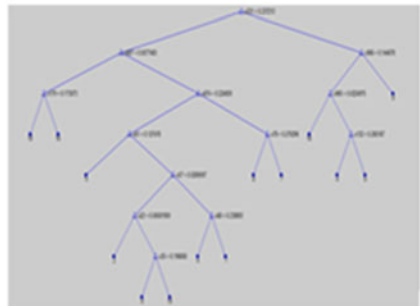**Fig. 14** Conversions to decimal

Decision Tree 1
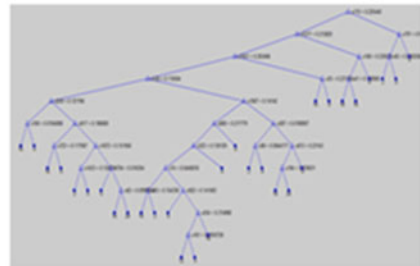
Decision Tree 2

Decision Tree 3

Decision Tree 4

Fig. 15 Training phase

4 Experimental Results

Our contribution is built on the image multi-label classification domain. For this purpose, we use scene dataset. It is a benchmark used for this purpose for several state-of-the-art algorithms [4]. It consists of 2000 images belonging to five natural scenes: mountains, desert, sunset, trees, and sea. We split it into 1600 training examples and 400 testing examples. Therefore, five evaluation metrics are used: Hamming Loss (HL), Ranking Loss (RL), One Error (OE), Average Precision (AP) and Coverage [4, 12]. These metrics are commonly used to evaluate the performance of multi-label classification, taking into consideration:

- The mis-classification of examples–labels pairs (HL)
- The order of the proper label (RL and OE)
- Proper label ranked above particular label (AP)

It is clear from the above sections that there are many important parameters to be set up. We will discuss them in each phase.

Table 1 shows the values of parameters taken in the experiments. The size of the image 128×128 in the first phase and the couple (scale, orientation) is used in several references [8]. The parameter K should be small to have enough labels during the training phase.

The last parameter is the single-label classifier used in LPP. We used the decision tree as classifier. It is a powerful nonparametric method independent of the distribution of the feature vector space.

Table 2 presents:

- Better results compared to the five evaluation metrics (HL, RL, AP, OE, Coverage) according to the major MIML methods found in the literature. (MIMLBoost, MIMLSVMmi, and MIMLNN) [4].
- The results of our method using four parameters (Size of the image, Scale, Orientation, and the number of clusters K).

The analysis of this table shows a significant enhancement in all metrics using our method compared with the others of MIML. The transformation of multi-label to single-label gives better accuracy (single label metric) using LPP. This affects positively the results in all multi-label metrics.

Table 1 Parameters for each phase

Phase	Parameters	Values
Preprocessing phase	The size of the image	128×128 64×64
Feature extraction	The scale, the orientation	(4,6) and (5,8)
K-Means	K	2, 3 and 4
LPP	Classifier	Tree decision

Table 2 Comparison between our method with the main state-of-the-art results

	HL↓	RL↓	AP↑	OE↓	Coverage↓
The best results (Algorithms[4])	0.185±0.08 (MIMLNN)	0.178±0.011 (MIMLBoost)	0.783±0.011 (MIMLSVMmi)	0.317±0.018 (MIMLSVMmi)	0.984.178 ±0.049 (MIMLBoost)
Size of the image is 128 × 128					
Scale_Orientation					
	K = 4				
(4, 6)	**0.0626**	**0.1212**	**0.9256**	0.0966	**0.5868**
(5, 8)	0.0727	0.1289	0.9222	**0.0905**	0.6379
	K = 3				
(5, 8)	**0.0589**	**0.1137**	**0.9325**	**0.0821**	**0.5693**
	K = 2				
(4, 6)	**0.0557**	**0.1089**	**0.9325**	**0.0795**	0.6073
(5, 8)	0.0615	0.1194	0.9272	0.0969	**0.5774**
Size of the image is 64 × 64					
Scale_Orientation					
	K = 4				
(4, 6)	**0.0606**	**0.1109**	**0.9323**	**0.0865**	**0.4861**
(5, 8)	0.0728	0.1341	0.9176	0.1118	0.5700
	K = 3				
(4, 6)	0.0726	0.1330	0.9187	0.1011	0.5612
(5, 8)	**0.0614**	**0.1114**	**0.9339**	**0.0867**	**0.5022**
	K = 2				
(4, 6)	**0.0604**	**0.1124**	**0.9337**	0.0846	**0.5090**

5 Conclusion

The aim of this chapter was to introduce a new framework for image multi-label classification. It is an improvement of the MIML framework. We presented the advantages of our method over three main MIML methods. The strengths of our method were found in its simplicity with regard to its implementation, solving the challenge of the description of the elementary characteristics from the image, and the correlation between labels and overall competitiveness in terms of the five evaluation metrics used. In the future, a new method can be developed in the feature extraction phase that optimizes the choice of each parameter.

References

1. R.S. Cabral, F. Torre, J.P. Costeira and A. Bernardino, Matrix completion for multi-label image classification. In Advances in Neural Information Processing Systems, pp. 190–198, 2011.
2. S.J. Huang., W. Gao and Z.H. Zhou, "Fast multi-instance multi-label learning". In Twenty-Eighth AAAI Conference on Artificial Intelligence, 21 June 2014.
3. Z. J. Zha, X. S. Hua, T. Mei, J. Wang,,G. J. Qi and Z. Wang, "Joint multi-label multi-instance learning for image classification". In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on (pp. 1–8). IEEE, June 2008.
4. ML. Zhang and ZH. Zhou, "A review on multi-label learning algorithms." IEEE transactions on knowledge and data engineering 26.8: pp. 1819–1837, August 2014.
5. M.R. Boutell, J. Luo, X. Shen, and C.M. Brown, "Learning multi-label scene classification". Pattern recognition, 37(9), pp. 1757–1771, 2004.
6. Z. Abdallah, A. El-Zaart, and M. Oueidat. "An Improvement of Label PowerSet Method Based on Priority Label Transformation." International Journal of Applied Engineering Research 11.16: pp. 9079–9087, 2016.
7. M. Nosrati, R. Karimi, M. Hariri and K. Malekian, "Edge detection techniques in processing digital images: investigation of canny algorithm and gabor method". World Applied Programming, pp. 116–21, March 2013.
8. S. Khan, M. Hussain, H. Aboalsamh, H. Mathkour, G. Bebis and M. Zakariah, "Optimized Gabor features for mass classification in mammography". Applied Soft Computing, pp. 267–80, 31 July 2016.
9. https://www.youtube.com/watch?v=-NZakhhB_Do
10. N. Sonawane and BD. Phulpagar, "Review on Content-Aware Image Re-sizing Using Improved Seam Carving and Frequency Domain Analysis", 2015
11. M. Kaur and U. Kaur, "Comparison between k-means and hierarchical algorithm using query redirection". International Journal of Advanced Research in Computer Science and Software Engineering, July 2013.
12. Abdallah, Ziad, Ali El Zaart, and Mohamad Oueidat. "Experimental analysis and comparison of multilabel problem transformation methods for multimedia domain." Applied Research in Computer Science and Engineering (ICAR), 2015 International Conference on. IEEE, 2015