

# An Improvement of Multi-Label Image Classification Method Based on Histogram of Oriented Gradient

Ziad Abdallah, Mohamad Oueidat, Ali El-Zaart

**Abstract**—Image Multi-label Classification (IMC) assigns a label or a set of labels to an image. The big demand for image annotation and archiving in the web attracts the researchers to develop many algorithms for this application domain. The existing techniques for IMC have two drawbacks: The description of the elementary characteristics from the image and the correlation between labels are not taken into account. In this paper, we present an algorithm (MIML-HOGLPP), which simultaneously handles these limitations. The algorithm uses the histogram of gradients as feature descriptor. It applies the Label Priority Power-set as multi-label transformation to solve the problem of label correlation. The experiment shows that the results of MIML-HOGLPP are better in terms of some of the evaluation metrics comparing with the two existing techniques.

**Keywords**—Data mining, information retrieval system, multi-label, problem transformation, histogram of gradients.

## I. INTRODUCTION

THE main objective of object recognition is to find a desired object in a collection of images or digital video. This identification process is a great challenge today in Computer Vision [1]. For each object in an image, there are interesting characteristics called features which distinguish it from other objects [2]. These features can be extracted to provide descriptive and distinctive characteristics of the object. They can be divided into two groups: a) Low-level visual characteristics or physical features [4]: Describe basic visual features such as the shape, the color, the texture, the forms, the movement and the regions, and allow us to find a connection between the pixels contained in a digital image and what people remember once they have observed an image and b) High-level visual characteristics or features logical based on the recognition of objects: provide information about objects and events that are appearing on the scene and the relationship that exists between them [3]. The purpose of this paper is to classify an image according to its context/content and identify the categories or classes that the object belongs (Image Classification). The objective is not to recognize a particular object, but check if the latter is present in the image and belongs to a certain number of Labels. A new type of machine learning framework proposed recently, named MIML

[4], where an object is described with many instances and can be assigned to multiple labels as shown in Fig. 1. For example, in the case of image annotation, the image contains different regions. These regions can be expressed as different examples called feature vector. At the same time, the image may be classified simultaneously for more than one category; in text categorization, a document can have different chapters or sections. It can belong to more than one category (scientific, religious, politics, etc.). MIML has been successfully applied to image text classification, image annotation, video annotation, ecological protection, and other tasks [4]-[7]. In this framework, an image is described with many instances and can be assigned to multiple labels. The MIML is a single label learning transformation. There are two ways to do this transformation: the first one transforms MIML to Multi-Instance Single-Label Learning (MISL) and then single-Instance Single-Label Learning (SISL). The second one transforms MIML to single-Instance Multi-Label Learning (SIML) and then SISL. Two most important existing techniques are proposed for this transformation which are (MIML-BOOST and MIML-SVM). Two limitations for these existing methods: they did not take into considerations the description of the elementary characteristics from the image [8] and the correlation between labels [4].

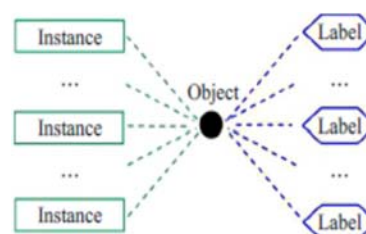


Fig. 1 MIML [4]

Our contribution is an algorithm, called MIML-HOGLPP, which handles the first limitation by using the histogram of gradients as feature descriptor and the second one by applying LPP as Multi-label classification method. The idea extracts the feature from image using Histogram of Oriented Gradient (HOG) algorithm and it is a feature extraction algorithm that takes into consideration the local representation, shape, and geometry of an image [13]. We apply then K-mean to cluster the image into similar groups. The final step was the learning phase using supervised learning algorithm Label Priority Powerset that transforms MIML problem to single label classification [5]. Each step of our new contribution will be described in Section III. The results in the experiment of five important evaluation metrics of multi-label classification show

Ziad Abdallah is with the Department of Mathematics and Computer Sciences, Beirut Arab University, Debbieh, Lebanon (phone: 961-3-840270; e-mail: z.abdallah@bau.edu.lb).

Mohamad Oueidat is with the Department of Maintenance and Industrial Engineering Lebanese University – University Institute of Technology, Saida Lebanon (e-mail: mohoueidat@yahoo.com).

Ali El-Zaart is with the Department of Mathematics and Computer Sciences, Beirut Arab University, Debbieh, Lebanon (phone: 961-70-870956; e-mail: elzaart@bau.edu.lb).

that our proposed method is competitive with the existing techniques of the literature. The remainder of this paper is organized as follows: Section II defines the tasks of MIML. Section III defines our proposed method. Section IV describes the experimental setup, and discusses the experimental results. Finally, the conclusions are given in Section V.

## II. RELATED WORK

### A. Multi- Label Classification

The traditional techniques of machine learning classification are designed for problems where each element to be classified, called instances, can only have one class [6],[7]. In reality, there are problems of classification where a single instance can belong to different disjoint classes, called labels.

In multi-label learning, there are two major types of algorithms: the transformation and the adaptation [5]-[8]. The transformation transforms the multi-label in one or more multi-class learning, which is resolved by traditional classification algorithms. One of the most popular is Binary Relevance (BR), which generates a binary classifier for each label, treating them independently. Label Powerset (LP) transforms the problem multi-label in a multi-class problem, creating a new class for each combination of different label that appears in the dataset [5]-[8].

The adaptation of algorithms is to extend classical algorithms to work directly in multi-label classification. They have adapted all types of classification techniques, for example, decision trees, neural networks, or instance-based algorithms [9].

Formally, an instance  $i$  is represented by a vector  $X_i = (x_1, x_2, \dots, x_m) \in \chi$  where  $\chi$  is the space of possible instances, and the labels that are associated with an instance are a subset  $Y_i \subseteq L$  (The set of all possible labels), with  $L = \{y_1, y_2, \dots, y_n\}$ . In this context, the problem of multi-classification labels is to find a function  $f: \chi \rightarrow 2^L$  from a set of training instances  $D = \{(X_j, Y_j) | 1 \leq j \leq p, \text{ where } p \text{ is the size of } \chi\}$ . For a new instance  $t$ , the function  $f(X_t)$  predicts a subset of labels  $Y_t \subseteq L$  [10].

### B. Multi-Label Image Classification

The amount of database of digital images has grown in a surprising way in the past few years. Thus, this situation demands efficiency in the search and retrieving methods for the extraction of images. The multi-labels image classification is the method that associates a label or a set of labels to an image. This automatic classification approach requires a set of images previously labeled and used to train a classifier. Then, the classifier is used to label the rest of the unlabeled images. A clear example of multi-label classification is the image annotation since each image usually contains more than one object of interest. The MIML is an algorithm that transforms the multi-label image to a single label classification. There are two ways to do this: The first one transforms MIML to MISL and then SISL. The second one consists to transform MIML to SIML and then SISL. A good review can be found in [4]. Fig. 2 shows an example of multi-label image classification, where we have multiple instances (two boats) and multiple labels

(three labels: sea, sunset, cloud).



Fig. 2 Multi-label image classification

### C. Feature Extraction Using HOG

This method, introduced for the first time by Dalal and Triggs [1], is based on the evaluation of histograms calculated on the basis of the direction and intensity of the gradients of the input image. Each histogram is obtained from a portion of the image called block, and by concatenating all blocks of an identification window, we get a descriptor of this, as shown in Fig. 3.

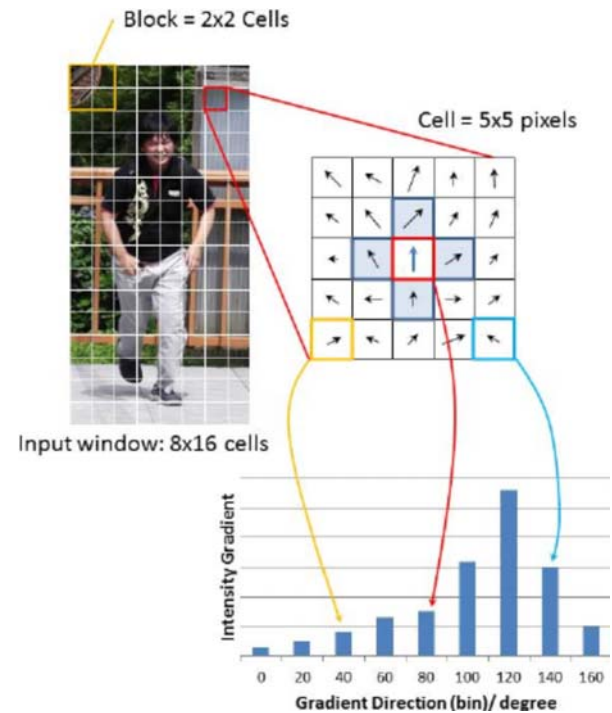


Fig. 3 Histogram of Oriented Gradient [12]

There are many ways to obtain the gradient of an image, but the simplest and the faster is the one which allows obtaining the best results [1]. It uses a simple mask 1-D:

$$h_x = [1, 0, -1]$$

$$h_y = [1, 0, -1]^T$$

Given an image (Img), we calculate the derivative with respect to the x axis and the y axis as:

$$dx = \text{Img} * h_x$$

$$dy = \text{Img} * h_y$$

and, for each pixel of the image, we obtain:

$$\text{Amplitude: } M(x,y) = \sqrt{dy^2 + dx^2}$$

$$\text{angle: } O(x,y) = \tan^{-1}\left(\frac{dy}{dx}\right)$$

A good literature review can be found in [1].

#### D.K-Means

The K-means algorithm [11],[14] is a clustering algorithm dividing groups of objects in K partitions on the basis of their attributes. It is efficient in managing large amounts of data and often terminates with an excellent local. The objective of the algorithm is to minimize the total variance intra-cluster (or the standard deviation). Each cluster is identified by a centroid or mid-point. The algorithm follows an iterative procedure:

- Initially creates K partitions and assigns to each partition the input points either randomly or using some heuristic information;
- Calculates the centroid of each group;
- Builds a new partition by associating each point of entry to the cluster whose centroid is closest to;
- Recalculates the centroids for the new cluster and so on, the centroids no longer move

### III. METHOD

This article proposes a new method that improves the MIML. The drawbacks of the existing methods of multi-label image classification did not take into consideration the following:

- a) The description of the elementary characteristics from the image: color, shape, regions, textures, and motion are some elementary characteristics.
- b) The correlation between labels: in multi-label learning, labels are correlated. For example, the label trees and desert are correlated to the yellow.

The structure of the algorithm, which addresses these drawbacks, can be generalized in three phases:

1. Preprocessing: the input image is optimized and adapted to meet the specifications of the algorithm and the feature extraction step.
2. Extraction of feature using HOG: the histogram of gradient is a feature extraction method that divides an image into small connected region called cells, where each pixel in the cell is compiled in order to create a gradient direction. The output of this step is a HOG feature vector. The advantage of HOG is that it takes into consideration the local representation, the shape, and the geometry of an image. The limitation of HOG is that the number of features is dynamic depending on the size of each block (set of cells).

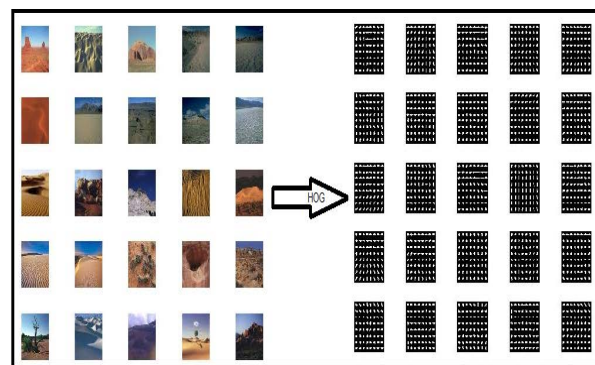


Fig. 4 Feature extraction using HOG method

3. 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. K-Means consists of grouping similar images into different k mutually exclusive clusters. The output of this step is K clusters C1, C2, ..., Ck. An image may belong to exactly one of these clusters. When K is small, the advantage of K-Means is better than Hierarchical Clustering in case of big amount of data. The disadvantage of K-Means is the difficulty in predicting the K which represents the number of clusters.
4. Label Priority Powerset transformation: we will use in this step, the transformation problem through breaking down the multi-label dataset into a single label dataset using the Label Priority Powerset transformation. The output of this step is a dataset  $D_s = \{(X_1, y_1), \dots, (X_p, y_p)\}$ , where  $X_i$  is the feature extracted from HOG method and  $y_i$  is the decimal conversion of binary multi-label after sorting the label by their importance. The importance of this step is the reduction of the complexity of learning process [5].
5. Classification: 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 that it applies single label classification in a multi-label problem. In case of K is big, the drawback is the complexity of the algorithm.

In brief, MIML-HOG extracts the important feature from image using Histogram HOG algorithm, solving the first limitation. Second, we apply the clustering technique K-means to group the similar images together into homogenous groups. This step was important for the classification process. Finally, we applied in the learning phase the supervised learning algorithm Label Priority Powerset that transforms MIML problem to single label classification, solving the second limitation.

### IV. EXPERIMENT

The purpose of the experiment compares MIMLHOG with the best results found in state-of-the-art algorithms of multi-label image classification. Therefore, five evaluation metrics



are used: HL, RL, OE, AP, and Coverage.

Our contribution is to build on the IMC 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.

We used MATLAB to develop our algorithm which is a powerful tool for research.

TABLE I  
COMPARISON OF MIMLHOG WITH THE BEST RESULTS FOUND IN LITERATURE

	HL	RL	AP	OE	Coverage
	.178	.167	.794	.299	.934
The best results (Algorithms[4])	(MIML-NN)	(MIML-Boost)	(MIML-SVMmi)	(MIML-SVMmi)	(MIML-Boost)
MIMLHOG	0.09	0.125	0.91	0.1	.7

The results of Table I prove that our algorithm is better in all metrics. Several reasons justify these results:

- Extracting the important features from images such as the orientation and the magnitude
- Clustering of images into homogenous groups which facilitates the learning in each cluster
- The average precision of MIML-HOG is high (0.91), affecting significantly the other metrics

#### V. CONCLUSION

This article deals with the classification of images from a learning perspective with multiple labels. It assesses the two representations that have been proposed in MIML and is compared with their efficiency by using HOG and LPP. Experimental results of the MIML-HOG with the state of art algorithms confirm that our algorithm is more appropriate for the multi-label image classification. The extraction of the important features such as the orientation and magnitude from image, the clustering of images into homogenous groups, and the learning in each cluster improves significantly the main metrics used for multi-label problem. We can conclude that further progress in this area is justified and could optimize the resolution of this problem.

#### REFERENCES

- [1] MZ Kurian, "Various Object Recognition Techniques for Computer Vision" Journal of Analysis and Computation 2011, vol. 7, pp. 39-47.
- [2] Y. Ramadevi, T. Sridevi, B. Poornima, and B. Kalyani. "Segmentation and object recognition using edge detection techniques" International Journal of Computer Science & Information Technology (IJCSIT), vol. 2, 2010, pp. 153-161.
- [3] T. Sikora. "The MPEG-7 visual standard for content description-an overview" IEEE Transactions on circuits and systems for video technology, vol. 11, 2001, pp. 696-702.
- [4] Z.H. Zhou, Min-Ling Zhang, Sheng-Jun Huang, and Yu-Feng Li. "Multi-instance multi-label learning" Artificial Intelligence vol. 176,2012, pp. 2291-2320.
- [5] Z. Abdallah, Ali El-Zaart, and Mohamad Oueidat, "An Improvement of Label PowerSet Method Based on Priority Label Transformation" International Journal of Applied Engineering Research, vol. 11, 2016, pp. 9079-9087
- [6] I. Pillai, Giorgio Fumera, and Fabio Roli. "Designing multi-label classifiers that maximize F measures: State of the art" Pattern Recognition, vol. 61, 2017, pp. 394-404.

- [7] N. Spolaôr, Everton AlvaresCherman, Maria Carolina Monard, and Hueli Diana Lee. "A comparison of multi-label feature selection methods using the problem transformation approach" Electronic Notes in Theoretical Computer Science, vol. 292, 2013, pp. 135-151.
- [8] M.R. Boutell, Luo, J., Shen, X. and Brown, C.M., "Learning multi-label scene classification" Pattern recognition, vol. 37, 2004, pp.1757-1771.
- [9] G. Tsoumakas, K. Ioannis Katakis, and Ioannis Vlahavas. "Data mining and knowledge discovery handbook." Mining multi-label data 2010.
- [10] M. Zhang and Z. Zhou, "A review on multi-label learning algorithms" Knowledge and Data Engineering, IEEE Transactions on, vol. InPress, 2013.
- [11] J. MacQueen. "Some methods for classification and analysis of multivariate observations" In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, vol. 1, 1967, pp. 281-297.
- [12] M. Hatto. "Acceleration of Pedestrian Detection System using Hardware-Software Co-design" 2015, pp 7.
- [13] H. Al-Shamlan and A. El-Zaart. "Feature extraction values for breast cancer mammography images" Bioinformatics and Biomedical Technology (ICBBT), 2010 International Conference.
- [14] A. Gumaci, A. El-Zaart, M. Hussien and M. Berbar, "Breast segmentation using k-means algorithm with a mixture of gamma distributions", Broadband Networks and Fast Internet (RELABIRA), 2012 Symposium.

**Ziad Abdallah** is born in CHHIM-Lebanon in 1972. He received his B.Sc. in Computer Sciences in 1994 from the Lebanese University, Lebanon. He granted his engineering in statistics from the Ecole Nationale de la Statistique et de l'Analyse de l'Information, France. He received his Master Recherche in modeling in the Agence universitaire de la Francophonie from Université Al Manar, Tunisia - Université Libanaise, Lebanon - Université de Reims, France - Université de Rennes I - France. At present (2015), he is applying his PhD studies in Information Technology at the BAU.

Ziad is working at the central administration of statistics as the head of IT department since 1998. He is also tutor in the Lebanese University since 1998, Beirut Arab University (BAU) since 2013, and ENA-Lebanon since 2009 for many courses related to Statistics, Computer Sciences and Information Systems. Ziad is responsible of the Consumer Price Index (CPI) project in Lebanon (2013-present).

He was the National coordinator in MEDSTAT Programme (Programme to enhance cooperation and statistical capacity building between EU and the mediterranean countries) of: MedENV (Environnement statistics), MedIS (Information system).

**Mohamad Oueidat** received the Degree in electrical engineering from the Electricity High School (Supélec), Paris, France, in 1993, and the D.E.A. (M.S.) degree in electrical engineering and the Ph.D. degree from the University of Paris VI, Paris, in 1993 and 1997, respectively. From 1997 to 2000, he was an Associate Professor with the Institute of Applied and Economic Sciences (associated with CNAM, Paris). From 2000 to 2014, he was an Associate Professor with the University Institute of Technology, Lebanese University, Beirut, Lebanon, where he is currently a Professor. His current research interests include power electronics, signal processing, data compression, and data mining.

**Ali El-Zaart** was a senior software developer at Department of Research and Development, Semiconductor Insight, Ottawa, Canada during 2000-2001. On 2001, he started his work in the capacity of assistant professor at the Department of Biomedical Technology, College of Applied Medical Sciences, King Saud University, Saudi Arabia. Afterward, on 2004, he transferred to the Department of Computer Science, College of Computer and Information Sciences. On 2010, he has been promoted to associate professor rank in computer science. Since 2011, he has been a faculty member at the Department of Mathematics and Computer Science, Faculty of Science, Beirut Arab University. On March 2016, he has been promoted to full professor rank in computer science. He received a M.Sc. (1996) and Ph.D. (2001) degrees in computer science from the University of Sherbrooke, Sherbrooke, Canada.

Professor Ali El-Zaart has published numerous articles and proceedings in the areas of image processing, remote sensing, computer vision, data mining, and smart cities. Her research interests are cancer detection in medical images Multilabel classification, and smart cities.