International Journal of Modern Trends in Engineering and Research
www.ijmter.com
e-ISSN: 2349-9745

A Review: MIML Framework and Image Annotation

Keyur Tank[1], Prof. Praveen Bhanodia [2], Prof. Pritesh Jain[3]
[1]*Computer Science & Engineering, PCST, Indore, India,keyur.kt@gmail.com*
[2] Asst. Professor& Head *Computer Science & Engineering, PCST, Indore,India,pcst.praveen@gmail.com*
[3] Asst. Professor *Computer Science & Engineering, PCST, Indore, India, pritesh.arihant@gmail.com*

Abstract— This review paper creates a bridge between MIML classification framework and Image annotation. There are generally four classification frameworks, known as Single Instance Single Label (SISL), Multi-Instance Learning (MIL), Multi-Label Learning (MLL) and Multi-Instance Multi-Label Learning (MIML). This paper introduces various classification frameworks with examples and related algorithms. An annotation is one type of metadata that can be attached to any video, image (2D/3D), text, audio and other data in the form of explanation, comments, navigation or presentational markup. This paper briefly introduces different types of annotation, annotation dataset, techniques and current research challenges in annotations

Keywords- MIML Classification Frameworks, Image Annotations, Image annotation.

I. INTRODUCTION

Nowadays databases contain lots of data and information which is impossible to analyze. To transformation this lots of amounts of data into useful information and knowledge we need some integration and classification techniques. Data mining provide many application fields such as marketing, engineering business, games, science, economics and bioinformatics. In today's knowledge-driven world use of multimedia or non-multimedia information produce enormous amounts of new information that we must process and aggregate to make it easier to understand. On the World Wide Web huge amount of high resolution images are being uploaded and retrieve every day. There are various technique and methods are available to deal with classification and annotations of such a dataset.

Classification is a machine learning technique in data mining which predict group membership for data instances and classes. Classification can be a Supervised or Unsupervised learning. In Supervised Classification the set of possible classes is known in advance. In Unsupervised Classification set of possible classes is not known. Current research challenges focuses on MIML classification framework which deals with multiple instances and multiple labels in any dataset. Annotation of dataset provides such a classification more efficiently. For example, YouTube video annotations are a new way to add interactive commentary to your videos. Classification separate data into learning (training) and classification (testing) sets. Training set is a dataset that is derived from original set and Testing set is a dataset that will be use to evaluate the performance of

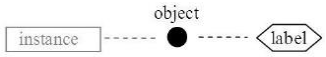
@IJMTER-2014, All rights Reserved

classifier or a model. Generally image contains multiple regions as a feature vector, so image annotation task is basically a MIML learning problem. New research in MIML classification framework deals with such a problem and generates annotation and learning methods more smoothly and accurately.

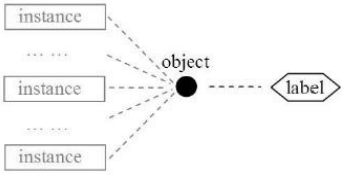
II. TYPES: FRAMEWORKS AND ANNOTATIONS

Classification frameworks can be building using instances or labels based on multivariate and/or univariate approaches.

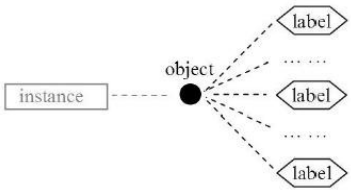
2.1 Types of Frameworks



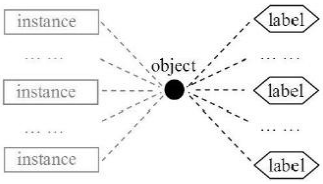
(a) Single Instance Single Label Learning (SISL)



(b) Multi-Instance learning (MIL)



(c) Multi-label learning (MLL)



(d) Multi-Instance Multi-Label learning (MIML)

Figure 1: Classification learning frameworks

In this traditional supervised learning classification method each instance of the dataset is associated with only one class.

2.1.2 Multi-Instance Learning (MIL)

In MIL [1, 2] classification method all the instance of the dataset is associated with single class. MIL can be use in classification of images, document or text categorization, drug activity prediction and activities related molecular.

2.1.3 Multi-Label Learning (MLL)

In MLL [3, 4] classification scheme each instance is associated with more than one class labels. Multi-label classification methods are increasingly required by modern applications,

@IJMTER-2014, All rights Reserved

such as video concept detection, text classification, weather forecast, gene functionality, music instrument recognition, semantic scene classification and music categorization.

2.1.4 Multi-Instance Multi-Label Learning (MIML)

MIML [5] learning is a new concept that consider the input and output ambiguities together. In MIML real-world objects are usually inherited with input ambiguity as well as output ambiguity. MIL and MLL are degeneration version of MIML shown in fig.

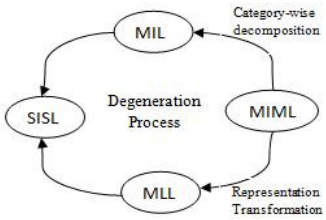


Figure 2: MIML Degeneration Process

2.2 Types of Annotation

Following are the different types of annotations which deal with the different types of datasets.

2.2.1 Text Annotation

Text can be non-formatted or with rich text formatting such as HTML markup. Text annotations should be in any text language and formats txt, pdf, rtf, Open Office, Word, etc or web page (html, xml). Main goal of a text annotation tool is to provide researchers how to find, create, and search media-rich annotations.

2.2.2 Image Annotation

It captures uploaded image or real-time image which is either 2D or 3D. It also provide image as an overlay or portion of the image. Image annotations should be possible in any web image formats like jpg, png, gif, svg, pdf.

2.2.3 Audio Annotation

It captures an uploaded file or a real-time audio recording. It provides time range (beginning and ending time) within audio and for the entire clip.

2.2.4 Video file annotation

It captures an uploaded clip or real-time video recording within a time range (beginning and ending time), video broadcasting, video as an overlay. Very few video players offer annotation features that are suitable for research, teaching or learning. Mostly such annotation are used primarily for marketing purposes, such as advertisements or for allowing social commentaries, banners etc. “Hug the world” is one of the best examples of video annotation given in YouTube.

III. LITERATURE REVIEW

3.1 MIML Classification Framework

MIMLBOOST [5] provide independent labels that decompose MIML task into a series of multi-instance learning tasks where all labels will be treat as a task. In the first step of MIMLBOOST, each MIML example is transformed into a set of number of multi-instance bags, where bag contains number of instances and labels. MIMLSVM [5] provide spatial distribution of the bags. Each bag provides relevant information for label discrimination

@IJMTER-2014, All rights Reserved

which measure distance between each bag and each representative bag identified using clustering methods. MIML-NN provides dependencies between different categories during decomposition into multiple set of classification problems using well-known Back-Propagation learning method (BP-MLL) [6]. Z.-H. Zhou et al. [7] Multi-instance multi-label learning with application to scene classification. Zhang et al. [8], also provide M3MIML: A Maximum Margin Method for Multi-Instance Multi-Label Learning. This method defines connection between instances and labels. In this method learning task is formulated as a quadratic programming (QP) drawback and implemented in its twin type.

3.2 Image Annotation

There are different techniques for the image annotation task. For image annotation, Jeon et al. [9] proposed the cross-media relevance model (CMRM) which tries to estimate the joint probability of the visual keywords and the annotation keywords on the training data set. This relevance model was further improved through continuous-space relevance model (CRM) [10], multiple Bernoulli relevance model (MBRM) [10], and dual cross media relevance model [11]. Carneiro et al. [12] proposed a probabilistic approach for this task. Guillaumin et al. [13] proposed a discriminatively trained nearest neighbor model in which tags of test images are predicted using a weighted nearest-neighbor model to exploit labeled training images. In [14], Zhang et al. introduced a regularization based feature selection algorithm to leverage both the sparsity and clustering properties of features, and incorporated it into the image annotation task.

3.3 MIML Framework and Image Annotation

T. Sumathi et al. provide survey on “Automatic Image Annotation and Retrieval using MIML Learning”, using different algorithms like MIMLBOOST, MIMLSVM, D-MIMLSVM, InsDif and SubCod algorithms. Cam-Tu Nguyen et al. proposed “Multi-Modal Image Annotation with Multi-Instance Multi-Label Latent Dirichlet Allocation (LDA)”. Z. H. Zhou et al. proposed the MIMLBOOST and MIMLSVM algorithms which achieve good performance in an application to scene (image) classification using MIML framework. Ameesh Makadia et al. [15] introduce a new baseline technique for image annotation that treats annotation as a retrieval problem.

IV. PROPOSED MIML FRAMEWORK AND IMAGE ANNOTATION

The MIML framework provides good performance against complicated objects with multiple semantic meanings under the MIML framework. MIML is more convenient and natural for representing complicated objects. Generally image contains multiple regions as a feature vector, so image annotation task is basically a MIML learning problem. MIML consider multiple instances and multiple labels together. Proposed framework deals with image dataset which reduces learning efficiency and consider indexing, browsing and retrieval of annotated image dataset from the database efficiently. Such a combination of MIML framework and image annotation can generates annotation and learning methods more smoothly and accurately. According to fig 4, Africa is a complicated high-level concept and the images belonging to Africa have great variance, thus it is not easy to classify it in class name Africa. Now, in fig 3 it is easy to define some low-level sub-concepts which are easier to learn with very less ambiguous, such as grassland, elephant, tree, lions etc. So it is easier to increase the concept Africa much easier than learning the concept Africa directly. In image classification the multiple labels of an image can be classify from different components (regions) in it. For example, in figure 2, there are three labels “sky,” “tiger,” and “grassland” is categorized by

@IJMTER-2014, All rights Reserved

three different regions, respectively, instead of the entire image and this same situation is shown in different framework. Image or text classification is a task of extracting information classes. Depending on the interaction between the analyst and the computer during classification, there are generally two types of classification named supervised and unsupervised. Supervised classification uses the spectral signatures obtained from training samples to classify an image. Unsupervised classification finds spectral classes (or clusters) in a multilabel and image without the analyst's intervention.

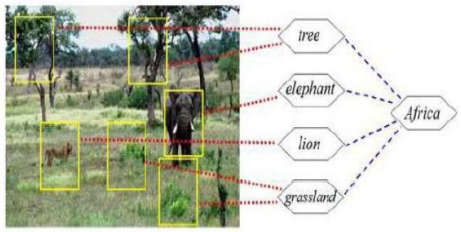


Figure 3: How Africa can be easier to learn through exploiting some sub-concepts.

Figure 4: Africa is a complicated high level concept

V. IMAGE ANNOTATION TECHNIQUES

In paper [16], different Image Annotation techniques are discussed as following. *Making use of Textual Information:* Most of images contain background information or text and associations can be used for image annotation. For example, search of BMW car will give all the images of different models while BMW 3 series will give only related images. *Manual Annotation:* This technique is best in terms of accuracy because it provides selection of keywords by the user but here it might be possible that user may forget annotation text later. *Image Annotation Based On Ontology:* Ontologies are structural framework or a set of concepts, and can be use in semantic web, biomedical, software engineering, library science etc. Given technique provide three layer architecture where bottom layer select features of images. These features are then mapped to semantically significant keywords in the middle layer. These keywords are then connected to schemas and ontologies on the top layer. *Semi Automatic Annotation:* This technique requires some user participants for some manual annotation process. It is very useful for dynamic databases but require User interfaces refinements to improve the feedback process. *Automatic Image Annotation:* This technique saves time by using image segmentation algorithms which divide images into different shapes. It uses the "global" features for automated image annotation. This modeling

@IJMTER-2014, All rights Reserved

framework is based on nonparametric density estimation, using the technique of "kernel smoothing". This technique is less reliable and produces more general annotation than manual annotation technique.

VI. RECENT PROGRESS

A good platform, architecture or framework always lead current research or work in progressive direction. By selecting and updating such a framework or architecture, goal towards the new research improve highly. As discussed current research challenges and progress are available in MIML framework. On this platform different types of dataset can be implementing and classify. During past days classification frameworks supported only single or multiple labels or instances, but not both together. Currently MIML framework deals with different learning environment, classifiers, algorithms, dataset, kernel methods, classification tools, feature extraction and description, classification evaluation metrics etc. To develop and train new algorithm, new MIML dataset are require shown in table 1. It shows different types of MIML dataset [17] with its number of classes, bags and instances.

Dataset	Classes	Dimension	Bags	Instances
MSRCv2	23	48	591	1,758
VOC 2012	20	48	1053	4,142
Birdsong	13	38	548	4,998
Carroll	26	16	166	717
Frost	26	16	144	565

Table 1: MIML dataset.

Microsoft Research Cambridge v2 (MSRCv2) and Visual Object Recognition Challenge (VOC 2012) both are image dataset. Bioacoustics Dataset (birdsong) is an audio dataset. Carroll and Frost both are artificial text based dataset. Table 2 shows the research in MIML framework and related different areas. This information was taken from the IEEE explorer between the years 2008 to 2014. It contains the details about year, author title and the description of that paper. This information will be helpful to understand latest research and different techniques in MIML framework. Furthermore, it also includes some of the topics from the ACM library. Other research is also available in this area where MIML perform most important part in classification.

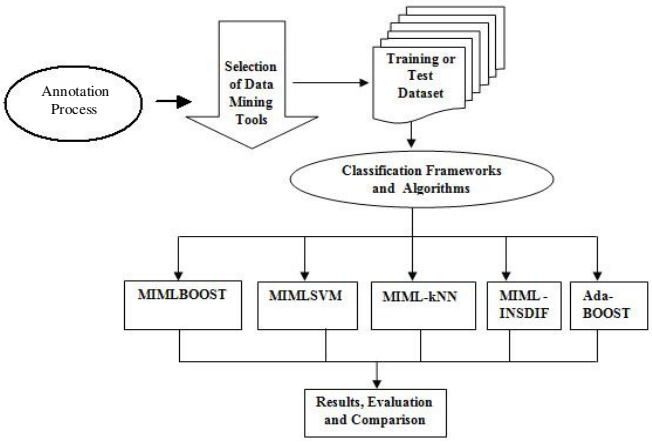


Figure 5: Methodology for MIML classification framework

Figure 5 shows the methodology to create MIML classification framework using annotation process. It also provides different algorithms and learning methods. At the end proposed model will be evaluate and compare using output result.

@IJMTER-2014, All rights Reserved

[1] Year	[2] Authors	[3] Title	[4] Description
[5] 2008	[6] Min-Ling Zhang ; Zhi-Hua Zhou	[7] M3MIML: A Maximum Margin Method for Multi-instance Multi-label Learning	[8] This work directly provides connection between labels and instances. This learning task uses quadratic programming to deal with different mathematical formulation. It provide linear model for every classes, where output of one class define the maximum prediction for other instances.
[9] 2009	[10] Shuangping Huang ; Jin, Lianwen	[11] A PLSA-Based Semantic Bag Generator with Application to Natural Scene Classification under Multi-instance Multi-label Learning Framework	[12] This mechanism generates set of instances from the image by using pLSA (Probabilistic Latent Semantic analysis) model.
[13] 2009	[14] Rong Jin ; Shijun Wang ; Zhi-Hua Zhou	[15] Learning a distance metric from multi-instance multi-label data	[16] This approach first estimate the association between instances of bags and class labels and then learn distance metric by a discriminative analysis method. This metric will be used to update association between instances and labels.
[17] 2009	[18] Lihua Guo, ; Jin, Lianwen	[19] The Generic Object Classification Based on MIML Machine Learning	[20] In this approach instances will be generated from the image and create a bag of instance. Then divide the image into four parts and calculate the histogram of edges of each four parts.
[21] 2009*	[22] Min-Ling Zhang and Zhi-Jian Wang	[23] MIMLRBF: RBF Neural Networks for Multi-Instance [24] Multi-Label Learning	[25] This approach proposed RBF neural networks for MIML framework. It also uses clustering and optimization methods to improve the performance.
[26] 2010	[27] Liang Peng ; Xinshun Xu ; Gang Wang	[28] An empirical study of automatic image annotation through Multi-Instance Multi-Label Learning	[29] This approach proposed ensemble method which evaluate four visual features and partition methods of two image
[30] 2010	[31] Min-Ling Zhang	[32] A k-NearestNeighbor Based Multi-Instance Multi-Label Learning Algorithm	[33] Proposedmethodusesk-nearestneighbor techniques. This method considers citers and its neighbors both together.
[34] 2010	[35] Nam Nguyen	[36] A New SVM Approach to Multi-instance Multi-label Learning	[37] This approach uses two optimization methods together. (1) A quadratic programming which reduces empirical risk and (2) An integer programming which create pair of single instance and label.
[38] 2011	[39] Jianjun Yan ; Qingwei Shen ; Jintao Ren ; Yiqin Wang ;	[40] A multi-instance multi-label learning approach to objective auscultation analysis of traditional Chinese medicine	[41] This method uses patients' speech including 5 vowels a, e, i, o, and u. Each patient in the dataset may have either one or both of the qi and yin deficiency syndromes. These syndromes will be considered as labels and all vowels in speech will be instance for any one label. After feature extraction data will be input to MIML classification.
[42] 2011	[43] Oksana Yakhnenko, Vasant Honavar	[44] Multi-Instance Multi-Label Learning for Image Classification with Large Vocabularies	[45] It proposed novel learning algorithm. Proposed algorithm deal with the discriminative multiple instance classifiers and provide the correlation among labels.
[46] 2011*	[47] Zhi-Hua Zhou [48] Xin-Shun Xu; [49] Xiangyang Xue; [50]	[51] Ensemble Multi-Instance Multi-Label Learning Approach [52] for Video Annotation Task	[53] It provides method of ensemble learning on video annotation dataset. It proposed En-MIMLSVM approach to deal with automatic annotation in video dataset.
[54] 2012*	[55] Zhi-Hua Zhou [56] Xin-Shun Xu; [57] Xiangyang Xue;	[58] Semi-Supervised Multi-Instance Multi-Label Learning for [59] Video Annotation Task	[60] It provides method of semi-supervised learning on video annotation dataset. It proposed semi-supervised MIML approach to deal with automatic annotation in video dataset.
[61] 2012	[62] Ying-Xin Li ; Shuiwang Ji ; Kumar, S. ;	[63] Drosophila Gene Expression Pattern Annotation through Multi-Instance Multi-Label Learning	[64] This approach deals with the patterns of Drosophila gene expression. It shows the manual annotation of images and then define terms and region in a group of images.
[65] 2013	[66] Qi Lou ; Raich, R. ; Briggs, F. ; Fern, X.Z.	[67] Novelty detection under multi-label multi-instance framework	[68] This method deals with novel-classes in a set of bags. The main goal is to determine whether the instances of bags are depended on a novel-classes or known classes.
[69] 2013	[70] Briggs, F. ; Fern, X.Z. ;	[71] Context-Aware MIML Instance Annotation	[72] This approach predicts instance labels instead of predicting label set of a bag. It uses ECC (ensemble

@IJMTER-2014, All rights Reserved

	Raich, R.		of classifier chains) through label correlation to improve instance level prediction.
[73] 2013	[74] Gang Zhang ; Xiangyang Su ; Yongjing Huang ;	[75] A sparse Bayesian multi-instance multi-label model for skin biopsy image analysis	[76] Proposed model deal with the complex relationship among features of local regions and annotation using set of 12700 skin biopsy images.
[77] 2014	[78] Wu, J. ; Huang, S. ; Zhou, Z.	[79] Genome-Wide Protein Function Prediction through Multi-instance Multi-label Learning	[80] This approach deals with the automated annotation of protein function.

TABLE 2: RESEARCH WORK ON MIML CLASSIFICATION FRAMEWORK BETWEEN YEAR 2008 AND 2014

*Paper in from ACM library

VII. CONCLUSION

This review paper provides an interface between MIML classification and annotation for advanced research. It describes a detail approach related MIML classification framework and Image annotation. It discusses different Frameworks, MIML algorithms, Annotation techniques, Classifiers, Learning methods and MIML Datasets. It provides one of the best solution and combination to deal with different classification and annotation problems. This paper also present latest research challenges and current status about MIML framework and annotation area. This review will be very helpful to the researchers who are new in MIML classification and Image annotation task.

REFERENCES

[1] Jun Yang, "Review of Multi-Instance Learning and Its applications", Carnegie Mellon University (CMU), pp 467-474, 2008

[2] Oded Maron, Tomás Lozano-Pérez, "A framework for multiple-instance learning", Proc. of the 1997 Conf. on Advances in Neural Information Processing Systems 10, pp.570-576, 1998.

[3] Grigorios Tsoumakas, Ioannis Katakis, and Ioannis Vlahavas. "Mining Multi-label Data". O. Maimon, L. Rokach (Ed.), Springer, 2nd edition, 2010.

[4] Mohammad S Sorower, "A Literature Survey on Algorithms for Multi-label learning", Computer Technologies and Information Sciences (CTIS), Oregon State University, OR 97330.December 2010.

[5] Zhi-Hua Zhou, Min-Ling Zhang, Sheng-Jun Huang, Yu-Feng Li, "Multi-Instance Multi-Label Learning". Nanjing University, Nanjing 210046, China. 2011.

[6] M.-L. Zhang and Z.-H. Zhou, "Multi-label neural networks with applications to functional genomics and text categorization," IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 10, pp. 1338–1351, 2006.

[7] Zhou, Z.-H., and Zhang, M.-L. 2007. "Multi-instance multi-label learning with application to scene classification", Proc. Advances in Neural Information Processing Systems 19, pp 1609-1616, 2007.

[8] Zhang and Zhou, "M3MIML: A Maximum Margin Method for Multi-Instance Multi-Label Learning", ICDM'08, 8th IEEE International conference on Data Mining, pp. 688-697, 2008.

[9] J. Jeon, V. Lavrenko, and R. Manmatha. "Automatic image annotation and retrieval using cross-media relevance models". In *SIGIR*, pp. 119–126, 2003.

[10] S. Feng, R. Manmatha, and V. Lavrenko. Multiple bernoulli relevance models for image and video annotation. In *CVPR*, pp. 1002–1009, 2004.

[11] J. Liu, B. Wang, M. Li, Z. Li, W.-Y. Ma, H. Lu, and S. Ma. "Dual cross-media relevance model for image annotation". In *MM*, pp 605–614, 2007.

[12] G. Carneiro, A. B. Chan, P. J. Moreno, and N. Vasconcelos. "Supervised learning of semantic classes for image annotation and retrieval". IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(3), pp 394–410, 2007.

@IJMTER-2014, All rights Reserved

[13] M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid. Tagprop: "Discriminative metric learning in nearest neighbor models for image auto-annotation". In *ICCV*, pp 309–316, 2009.

[14] S. Zhang, J. Huang, Y. Huang, Y. Yu, H. Li, and D. Metaxas. "Automatic image annotation using group sparsity". In *CVPR*, pp 3312–3319, 2010.

[15] Ameesh Makadia, Vladimir Pavlovic, and Sanjiv Kumar. "A New Baseline for Image Annotation", International Journal of Computer Vision 90(1), pp 88-105, 2010.

[16] Reena Pagare and Anita Shinde. "A Study on Image Annotation Techniques". In IJCA, pp 0975 – 8887, 2012.

[17] Forrest Briggs, Xiaoli Z. Fern, Raviv Raich, "Context-Aware MIML Instance Annotation", ICDM, IEEE, 2013.

@IJMTER-2014, All rights Reserved