International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 05 | May -2017 www.irjet.net p-ISSN: 2395-0072

A DEEP LEARNING MODEL FOR IMAGE CLASSIFICATION

Aparna R. Rout[1], Prof.(Dr.) Sahebrao B. Bagal[2]

[1]*PG Student, Department. Of Electronics and Telecommunication, Late G.N Sapkal College of Engineering, Nashik, Maharashtra, India*
[2]*Professor, Late G.N Sapkal College of Engineering, Nashik, Maharashtra, India*

---***---

Abstract ***In Image classification we classify image into one of the predefined classes. In conventional way, people use different computer vision techniques to extract features from images different machine learning algorithms uses these extracted features to classify the images. Various Supervised machine learning algorithms have been applied to multilabel image classification problems which have also brought successful results. It has become very difficult task to classify the images into interpretative classes. Apart from various learning algorithms the accuracy and performance of the model mostly depends on the trained dataset and the algorithm used. In this paper we have proposed a system to classify the scenery images into different groups of sunset, desert, mountains, trees and sea. In this paper the Current approaches for image classification make essential use of machine learning methods. We focus on deep learning techniques for feature extraction and classification of images. For multi class image classification we created dataset having landscapes scenes of sunset, desert, mountains, trees and sea. For multi label we use natural scenes dataset. In multi label classification problem an instance can have presence of more than one class among the given classes. There are methods to solve multi label classification problem but most of them are based on creating number of binary model equal to the number of labels and this technique is nothing but the Binary Relevance method. In this project, we propose a model which does not require creating multiple binary models instead it has single model which predicts the probabilities of different labels and uses probabilistic threshold values for respective label to convert those probabilities into presence and absence of class/label.***

Keyword: Supervised, Multilabel, Deep learning, machine learning, Binary Relevance Method

1. INTRODUCTION

The term image classification refers to the labelling of images into one of a number of predefined categories. Classification is a task to identify the class/category of new instance based on training set whose classes are known. In Image Classification, we classify an image into one of the predefined classes or multiple classes at the same time. With the rapid increase of digital photography, image understanding becomes increasingly important. Image semantic understanding is typically formulated as a *multi-class* or *multi-label* learning problem [1]. . In traditional supervised learning, an object is represented by an instance (or feature vector) and associated with a class label. Formally, let X denote the instance space (or feature space) and Y the set of class labels. Then the task is to learn a function $f : X \rightarrow Y$ from a given data set $\{(x1, y1), (x2, y2), \cdots, (xm, ym)\}$, where xi € X is an instance and yi € Y the known label of xi. Although the above formalization is prevailing and successful, there are many real-world problems which do not fit this framework well, where a real-world object may be associated with a number of instances and a number of labels simultaneously. For example, an image usually contains multiple patches each can be represented by an instance, while in image classification such an image can belong to several classes simultaneously, e.g. an image can belong to mountains as well as Africa [2]. Although this is always not a difficult task for humans, it has proved to be an extremely difficult problem for machines. Image classification is a widely studied problem in the field of Machine Learning for which there are many techniques and algorithms proposed. Deep Learning is one such technique. This work focuses on the application of deep learning algorithms for multi-label, multi-class Image Classification.

To summarize the proposed deep learning method using ConvNet for multi label image classification has the following key features compared to the existing methods:
1. A multi stage deep learning framework is designed to do the local discrimination and build local classifier.
2. Convolutional Neural Network is used for feature extraction and the Neural network is used for the classification purpose.
3. CNN is learned in a multi-instance learning fashion using the transfer learning

4. The Deep neural network thus receives the required features with the labels and classifies it accordingly.
5. One of deep learning's main advantages over all previous neural nets and other machine-learning algorithms is its capacity to extrapolate new features from a limited set of features contained in a training set. That is, it will search for and find other features that correlate to those already known.

1.1 Comprehensive Block Diagram

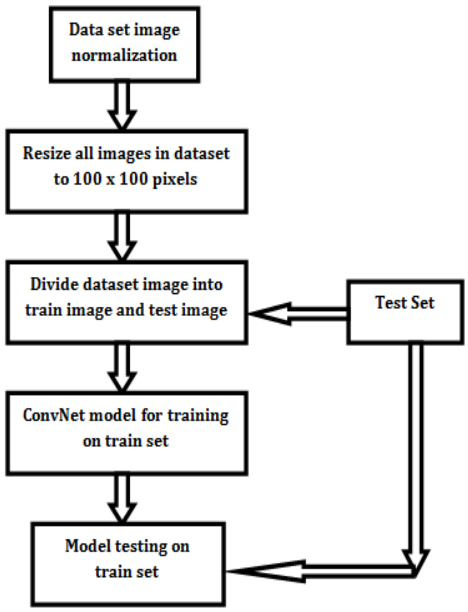

Fig1:-Block diagram of the heartwall segmentation

1.2 Components of the Block Diagram

Formation of dataset:

For Multi Class image Classification, we created the dataset by crawling five different products such as Trees, Sunset, water, desert and mountains. For each of the product we crawled 2000 [2] images and manually deleted those images which are not relevant to our task. Here each product represents a class. The "original" part contains 2000 natural scene images. The "processed" part contains data sets for multi-instance multi-label learning. This part is not big, about 618Kb (608Kb after compression). The image data set consists of 2,000 natural scene images, where a set of labels is artificially assigned to each image. The following table gives the detailed description of the number of images associated with different label sets, where all the possible class labels are desert, mountains, sea, sunset and trees. The number of images belonging to more than one class (e.g. sea + sunset) comprises over 22 of the data set, many combined classes (e.g. mountains + sunset + trees) are extremely rare. On average, each image is associated with 1.24 class labels. The Scene Classification Image data consists of 2000 [4] natural scene images, where a set of labels is artificially assigned to each image. The Table 4.1 gives a description of the number of images associated with different label sets, namely desert, mountains, water, sunset and trees.

Pre-processing:

First we resized all images in to 100x100 because CNN requires fixed size image as input. And split dataset into 80% and 20% where test set has 400 examples from each class and training set has 1600 images. Resized all images to 100x100 pixels and created two sets i.e. train set and test set. The labels for each image will be vector of one's and zero's for example [0, 1, 0, 1, 1] where 1 represent the presence and 0 represent the absence of a particular label. We have divided the dataset into train set and test set where train set contains 80% part of the whole dataset and rest part we treated as test set. After all epochs we got 90% train accuracy and 86% test accuracy.

ConvNet model:

Convolutional Neural Networks are very similar to ordinary Neural Networks. They are made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. ConvNet architectures make the explicit assumption that the inputs are images, which allows us to encode certain properties into the architecture. These then make the forward function more efficient to implement and vastly reduce the amount of parameters in the network.

2.MULTI-LABEL IMAGE CLASSIFICATION TECHNIQUES

Related research on image classification can be summarized along three paradigms: Multi Label Batch mode active learning [11] Binary Relevance Method [6] Multi-Label k-Nearest Neighbour (ML k-NN) [4], Random-Label set (RAkEL) [5]

2.1 Multi-Label Batch mode active learning:

One of the principal difficulties in applying supervised learning techniques to image classification problems is the large amount of labeled training images that are required. In many cases, unlabeled images are easy to obtain, while annotation is expensive or time consuming. This necessitates active learning [12], [13] which allows the learning algorithm to actively select the images from which it learns. Its key idea is to find the most informative images for annotation with respect to the maximal improvement to current classifier's performance, thereby reducing the annotation cost. Active learning is performed in an iterative fashion. Taking traditional binary myopic active learning as an example, each its learning iteration selects the example with the highest informativeness score for annotation, and the classifier is retrained on the training dataset with new labeled example included. The learning process continues until all the annotation resources are depleted or the obtained classifier's performance is accurate as desired.

In order to overcome some drawback, batch mode active learning has recently attracted increasing attentions. It aims to select a batch of informative examples instead of a single example for annotation at each learning iteration. The key difficulty comes from the potential information overlapping within the selected examples at each iteration, namely the selected examples need to be not only informative but also diverse. In this paper, Bang Zhang tackles image classification as a multi-label batch mode active learning (MLBAL) problem.

The proposed method in this paper is developed with the consideration of all these issues:

(1) A score function is defined to measure the informativeness of example-label pairs. It is designed based on both likelihood maximization (on labeled data) and uncertainty minimization (on unlabeled data).

(2) In order to take advantage of informative label correlations, we define cross-label uncertainty which gauges the disagreement between the mined label correlation and the label co-occurrence possibility from the learned classification mode. Kullback-Leibler (KL) divergence [14] is utilized to measure such cross-label uncertainty.

(3) The proposed method considers not only pair-wise but also higher order label correlations. An auxiliary compositional label is defined as a combination of primary labels with the interest of utilizing informative high order correlations.

(4) For informative label correlation discovery, an efficient data mining method called association rule mining [15] is adopted. An informative correlation can be found from one association rule, and its informativeness is measured by the *support* and *confidence* of the rule.

The first experiment is conducted on the scene dataset which contains natural scene images with multiple labels. Images are collected from COREL image collection and Internet. Sample images are shown in Fig. 1. It is originally used in [16]. There are 5 class labels: *desert*, *mountains*, *sea*, *sunset* and *trees*. Adopting the method developed in [17], each image is depicted by a bag of nine 15-dimensional feature vectors. A subset of the whole dataset is used. It contains all the 457 multi-label images and 250 single-label images. The statistic about different labels can be found in Table I. For the proposed method, multiple-instance multiple label SVM [16] is used to generate prediction models. Specifically, images are depicted by adopting multi-instance learning framework [18]–[20]. Each image is represented as a collection of nine instances (image patches) by following the work in [17]. Then, as described in [16], a multi-instance multi-label SVM is used to generate classification models: Constructive clustering method [21] is applied first to convert multi-instance examples to standard single-instance examples. Multi-label SVM [7] is then used to generate multi-label classifiers. All the methods perform 3 iterations of learning (initially 400 example-label pairs, then each iteration queries 100 more example-label pairs until 700 example-label pairs). The final results are averaged over 5 times random runs. Five different evaluation metrics are used including Hamming loss, one-error, coverage, ranking loss, and average precision, as used in [22], [16].

Fig. 1. Some sample images from multi-label natural scene dataset.

Label	Image	Label	Image	Label	Image
D	50	D+SU	21	SU+T	28
M	50	D+T	20	D+M+SU	1
S	50	M+S	38	D+SU+T	3
SU	50	M+SU	19	M+S+T	6
T	50	M+T	106	M+SU+T	1
D+M	19	S+SU	172	S+SU+T	4
D+S	5	S+T	14		

Table 2.1 Statistics on scene dataset (D: desert, M:mountains, S:sea, SU: sunset, T:trees)

2.2 Binary Relevance Method:

Traditional single-label classification methods are concerned with learning from a set of examples that are associated with a single label y from a set of disjoint labels L, $|L| > 1$ [3]. However, there are several scenarios where each instance is labeled with more than one label at a time, *i.e.*, each instance $\mathbf{x}_i$ is associated with a subset of labels $Y_i \in L$. In this case, the classification task is called

multi-label classification.
Problem transformation methods map the multi-label learning task into one or more single-label learning tasks. When the problem is mapped into more than one single-label problem, the multi-label problem is decomposed into several independent binary classification problems, one for each label which participates in the multi-label problem. This method is called Binary Relevance (BR). Binary Relevance Method which converts multi-label problem to multi-class problem means if there are k labels so k binary models are made and each model will make a decision of presence or absence of that particular label with Linear SVM as a base classifier Table [2.2]. Table 2.2 shows the working of Binary Relevance Method where leftmost table shows that first classifier will build model by taking class 1 as first class and rest of the classes i.e.2; 3; 4&5 together as second class. The Binary Relevance method is a problem transformation strategy that decomposes a multi-label classification problem into several distinct single-label binary classification problems, one for each of the q labels in the set L = {y1, y2,.... yq}.

Ex.	Label	Ex.	Label	Ex.	Label
class 1	λ_1	class 1	$*\lambda_2$	class 1	$*\lambda_3$
class 2	$*\lambda_1$	class 2	λ_2	class 2	$*\lambda_3$
class 3	$*\lambda_1$	class 3	$*\lambda_2$	class 3	λ_3
class 4	$*\lambda_1$	class 4	$*\lambda_2$	class 4	$*\lambda_3$
class 5	$*\lambda_1$	class 5	$*\lambda_2$	class 5	$*\lambda_3$

Table 2.2 Transformed datasets produced by Binary Relevance (BR) Method

2.3 Multi-Label k-Nearest Neighbour (ML k-NN):

M Zhang and Z Zhou extended the famous k-Nearest Neighbour approach to Multi Label k-Nearest Neighbour (ML k-NN) which is a lazy learning approach it actually calculates the prior probabilities and conditional probabilities on k nearest instances and from these posterior probabilities is calculated for presence and absence of each label and make decision on it. [4]. In natural scene classification, each natural scene image may belong to several image types (classes) simultaneously. Through analyzing images with known label sets, a multi-label learning system will automatically predict the sets of labels for unseen images. The above process of semantic scene classification can be applied to many areas, such as content-based indexing and organization and content-sensitive image enhancement, etc. In this paper, the effectiveness of multi-label learning algorithms is also evaluated via this specific kind of multi-label learning problem. The experimental data set consists of 2000 natural scene images, where a set of labels is manually assigned to each image. The number of images belonging to more than one class (e.g. sea +sunset) comprises over 22% of the data set; many combined classes (e.g. mountains + sunset + trees) are extremely rare. On average, each image is associated with 1.24 class labels. In this paper, each image is represented by a feature vector using the same method employed in the literature [7]. Concretely, each color image is firstly converted to the CIE Luv space, which is a more perceptually uniform color space such that perceived color differences correspond closely to Euclidean distances in this color space. After that, the image is divided into 49 blocks using a 7×7 grid, where in each block the first and second moments (mean and variance) of each band are computed, corresponding to a low-resolution image and to computationally inexpensive texture features, respectively. Finally, each image is transformed into a 49×3×2=294-dimensional feature vector. [8]

2.4 Randomk-Labelset (RAkEL) :

G. Tsoumakas, L. Katakis and L. Vlahavas proposed an ensemble method called Randomk-Labelset (RAkEL) [5] method which is improvement on Label Powerset method by using this algorithm itself. The label power set is a straight forward method that considers each unique set of labels in a multi-label training data as one class in the new transformed data.[9] Therefore, the new transformed problem is a single label classification task. In RAkEL it makes new small distinct subsets of size m of labels (k labels) then from training dataset take examples which contains any one of the label which are in a particular subset and that new datasets to label Powerset method to classify like we can have maximum of (k/m) models so for a test example is passed through all those models and voting is done for final prediction.
It focuses on the label Powerset (LP) multilabel learning method [7], [8], which considers each subset of L, hitherto called *Labelset*, that exists in the training set as a different class value of a single-label classification task. LP is an interesting approach to study, as it has the advantage of taking label correlations into consideration. In order to deal with the aforementioned problems of LP, this work proposes randomly breaking the initial set of labels into a number of small-sized labelsets, and employing LP to train a corresponding multi-label classifier. This way, the resulting single-label classification tasks are computationally simpler and the distribution of their class values is less skewed. The proposed method is called RA*k*EL (RAndom *k* labELsets) [10], where *k* is a parameter that specifies the size of the labelsets. The main idea in this work is to randomly break a large set of labels into a number of small-sized labelsets, and for each of them train a multi-label classifier using the LP method. For the multi-label classification of an unlabeled instance, the decisions of all LP classifiers are gathered and combined. For simplicity, we only consider labelsets of the same size, *k*. A labelset $R \in L$ with $k = |R|$ is called *k*-labelset. Therefore, the proposed approach is dubbed RA*k*EL (RAndom *k* labELsets).Therefore; the new transformed problem is a single label classification task Table 2.3. In RAkEL it makes

new small distinct subsets of size m of labels (k labels) then from training dataset take examples which contains any one of the label which are in a particular subset and that new datasets to label Powerset method to classify like we can have maximum of $\binom{k}{m}$ models so for a test example is passed through all those models and voting is done for final prediction.

Table 2.3: Transformed data using Label Powerset method

Instances	Label Sets
1.	{λ1}
2.	{λ3, λ5}
3.	{λ2, λ4}
4.	{λ2,λ3,λ5}
5.	{λ1, λ4}

3. CONCLUSION

Thus here we have summarized different multi-label image classification techniques. Multi-label image classification using Multi label Batch Mode Active Learning (MLBAL) achieves better results than any other image classification technique reviewed. Our final conclusion is robustness of the algorithm which depends on highly specific deep learning algorithm and the neurons in the neural network.

REFERENCES

[1] Joint Multi-Label Multi-Instance Learning for Image Classification - Zheng-Jun Zha, Xian-Sheng Hua, Tao Mei, Jingdong Wan, Guo-Jun Qi, Zengfu Wang- Department of Automation ‡Internet Media Group University of Science and Technology of China ‡Microsoft Research Asia

[2] Multi Label Natural Scenes Dataset, http: // lamda. Nju. edu. cn/ data_ MIMLimage. Ashx

[3] Tom M. Mitchell. Machine Learning. McGraw-Hill Education, 1997.

[4] Zhang, M.L.; Zhou, Z.H. (2007), ML-KNN: A lazy learning approach to multi-label learning, Pattern Recognition 40 (7): 20382048.

[5] Grigorios Tsoumakas, Ioannis Katakis, Ioannis Vlahavas., Random k- Labelsets for Multi-label Classification, IEEE VOL. 23, NO. 7 2011

[6] Multi-label Problem Transformation Methods by Everton Alvares Cherman, Maria Carolina Monard, Jean Metz.

[7] M.R. Boutell, J. Luo, X. Shen, C.M. Brown, Learning multi-label scene classification, Pattern Recognition 37 (9) (2004) 1757–1771.

[8]GrigoriosTsoumakas, Ioannis Katakis, and Ioannis Vlahavas. Mining Multi-label Data

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

[10] G. Tsoumakas and I. Vlahavas, “Random k-labelsets: An ensemble method for multilabel classification,” in Proceedings of the 18th European Conference on Machine Learning (ECML 2007), Warsaw, Poland, September 17-21 2007, pp. 406–417.

[11] Multilabel Image Classification via High-Order Label Correlation Driven Active Learning Bang Zhang, Member, IEEE, Yang Wang, Senior Member, IEEE, and Fang Chen, Senior Member, IEEE

[12]D. Angluin, “Queries and concept learning,” Mach. Learn., vol. 2, no. 4, pp. 319–342, 1988.

[13] B. Settles, “Active learning literature survey,” Dept. Comput. Sci., Univ. Wisconsin–Madison, Madison, WI, USA, Tech. Rep. 1648, 2009

[14] S. Kullback and R. Leibler, “On information and sufficiency,” Ann. Math. Stat., vol. 22, no. 1, pp. 79–86, 1951.

[15] R. Agrawal, T. Imieli´nski, and A. Swami, “Mining association rules between sets of items in large databases,” ACM SIGMOD Rec., vol. 22, no. 2, pp. 207–216, 1993.

[16] Z. Zhou and M. Zhang, “Multi-instance multi-label learning with application to scene classification,” in Proc. 20th Annu. Conf. NIPS, 2006, pp. 1–8.

[17] O. Maron and A. Ratan, “Multiple-instance learning for natural scene classification,” in Proc. 15th ICML, 1998, pp. 341–349.

[18] T. G. Dietterich, R. H. Lathrop, and T. Lozano-Pérez, “Solving the multiple instance problem with axis-parallel rectangles,” Artif. Intell. vol. 89, no. 1, pp. 31–71, 1997.

[19] G. Herman, G. Ye, J. Xu, and B. Zhang, “Region-based image categorization with reduced feature set,” in Proc. IEEE 10th Workshop Multimedia Signal Process. Oct. 2008, pp. 586–591.

[20] B. Zhang, Y. Wang, and W. Wang, “Multiple-instance learning from multiple perspectives: Combining models for multiple-instance learning,” in Proc. IEEE WACV, Jan. 2012, pp. 481–487.

[21] Z.-H. Zhou and M.-L. Zhang, “Solving multi-instance problems with classifier ensemble based on constructive clustering,” Knowl. Inf. Syst., vol. 11, no. 2, pp. 155–170, 2007.

[22] R. Schapire and Y. Singer, “BoosTexter: A boosting-based system for text categorization,” Mach. Learn., vol. 39, nos. 2–3, pp. 135–168, 2000.