Engineering and Image Classification Framework Using Multi Instance Learning with KCCA Algorithm

P. Bhuvaneswari and S. Chitrakala
Easwari Engineering College, Anna University of Technology, Chennai, TamilNadu, India

Abstract: Image classification is a challenging task with many applications in computer vision. Images are annotated with multiple keywords that may or may not be correlated. Therefore, image classification may be naturally modelled as a Multiple Instance Learning problem. The main challenge of this problem is that usually classes are overlapped and correlated. In single label classification the correlation among instance is not taken into account. In an image the instance may belong to several classes. The correlations among different tags can significantly help predicting precise labels for improving the performance of multi label image classification. This study proposes a method Kernel Canonical Correlation Analysis (KCCA) and Multi Instance Learning for multi label image classification, for improving the performance of classification accuracy. The proposed framework comprises an input image which can be partitioned into image patches and features can be extracted. It breaks the original training set into several disjoint clusters of data. It then trains a multi-label classifier from the data of each cluster. The K means clustering is used to perform automatic instance cluster. Kernel canonical Correlation analysis can be made between disjoint clusters to know exact correspondence between image patches. Multi Instance Learning is one potential solution to address the issue of huge inter-concept visual similarity and improve the classification accuracy. The proposed approach reduces the training time of standard multi-label classification algorithms, particularly in the case of large number of labels.

Keywords: Clustering, correlation analysis, multiple instance learning, normalized

INTRODUCTION

Image data are typically Multi Labelled, each image belongs to several classes. In Single instance where classes are mutually exclusive and most approaches treat a learning task as multiple binary classification tasks. These way potential correlations among the labels are not taken into account. But in Multi-label where classes are not mutually exclusive and an instance that may belong to several classes simultaneously. The Single label classification is concerned with learning with a single label \( \lambda \) from a set of disjoint labels \( L, |L|>1 \). In multi-label classification, the examples are associated with a set of labels \( Y \subseteq L \).

However, in Multi Instance Learning approaches that attempt to take correlations into account. The major issues are huge inter concept similarity, huge intra concept diversity. Kernel canonical Correlation analysis can be made between disjoint clusters to know exact correspondence between image patches.

The main challenge of this problem is that usually classes are overlapped and correlated. The multi modal visual features can be extracted to solve the issue of inter concept similarity. It automatically clusters the image instance based on similarity and performs correlation analysis on those clusters to find exact correspondence between image patches. The inter related classifier is used to classify image instance based on domain specific class. It improves the performance of image classification.

The above problem has stimulated us to propose a Image Classification system to classify image based on correlation analysis and multi instance learning algorithm. It significantly improves the classification accuracy and reduces the training time of multi label classification algorithms.

LITERATURE REVIEW

The effective multi-label classification methods such as Multi-label boosting strategy which learning the correlation between image segments and the text tag in a set of training images. SVM based multi-label active learning model is explored for image classification (Kumar et al., 2011). Multi-Instance Learning (MIL) based image classification takes the relations between labels and regions into account. In this paradigm, an image is regarded as a bag consisting of multiple instances (i.e., regions). MIL allows of only labeling images at the Image level, instead of labeling at region level, when building classifiers. For a specific semantic
label, a bag is labeled positive if at least one instance has the corresponding semantic meaning; otherwise, it is negative (Chen and Wang, 2006). SVM based multi-label active learning model is explored for image classification. It considers the Correlation of words among the semantic keyword of images. It minimize the classification error are different due to the inherent label correlations (Nasierding and Kouzani, 2010). Multi label classification considers correlation among labels. Problem Transformation (PT) and Algorithm Adaptation (AA) to convert multiple binary classifications into single label classification. Adapt binary classification (Guo-Jun et al., 2010). The critical problem remained in the existing approaches is how to exploit the correlations between class labels. Multi Instance Learning with Kernel Canonical Analysis is proposed to identify correlation among image patches to classify images based on domain specific. It improves the image classification accuracy.

**METHODOLOGY**

**System architecture:** Each image region is treated as one instance and multimodal visual features are extracted. The region-based visual features are Colors, Edges and Textures (Contourlet). Contourlet transform improve the performance of edge detection in initial stage. An Automatic Instance Clustering and Kernel Canonical correlation analysis is performed for identifying exact correspondence between image instances. The proposed system introduces Multi Instance Learning with kernel canonical correlation analysis is used to know extract correspondence between image patches and SVM classifiers are used to classify images based on domain specific. The input image can be segmented into image patches.

The functional architecture for proposed system is given in Fig. 1. It has the following modules:

- Feature Extraction
- Multi Instance Learning
- Classifier

To avoid the issue of image segmentation for feature extraction, four grid resolutions are alternatively used for image partition and feature extraction. Each image region is treated as one instance and multimodal visual features are extracted from each image instance. The following region-based visual features extracted and they are colors, edges, Texture (Contourlet).

There are two well-accepted approaches for classifier training: discriminative approach and generative approach. The major weakness of the discriminative approach is that it cannot tackle the issue of huge inter-concept visual similarity effectively. The major weakness of the generative approach is that it cannot handle the high-dimensional multimodal visual features effectively. It is very attractive to integrate them to tackle the issues of huge intra-concept visual diversity and huge inter-concept visual similarity more effectively. To address these two challenging issues more effectively, a Multi Instance Learning is developed in this study to achieve more accurate training of classifiers.

**Feature extraction:** Each multi label image can be partitioned into set of image instances/image regions. Each image region is treated as one instance and multimodal visual features are extracted from each image instance to characterize its various visual properties more sufficiently. The following region-based visual features are extracted which are colours, shape, texture (contourlet). To capture the texture feature Contourlet Transform (CT) is used.
Contourlet feature extraction: Contourlet transform is a multi scale and directional image representation that uses first a wavelet like structure for edge detection. Contourlet transform can be divided into two main steps: Laplacian Pyramid (LP) decomposing and Directional Filter Banks (DFB) (Zheng-Jun et al., 2008).

Figure 2 shows sparse expansions for typical images having smooth contours. Each type of these visual features is used to characterize one certain type of the visual properties of the image instances and a suitable base kernel is designed for each type of visual features for characterizing one certain type of the visual similarity between the image instances.

Multi instance learning: The proposed system uses k-means algorithm for automatic instance clustering and kernel canonical correlation analysis to perform correlation analysis between those two clusters. The image Instances in the positive bags and the negative bags are first partitioned into multiple clusters according to their visual similarity contexts. Inter-cluster correlation analysis is further performed by kernel canonical correlation analysis for identifying the exact correspondences between multi labels and the image instances. Multi Instance learning is used to solve the issue of inter concept similarity. It improves the multi label image classification accuracy.

K-means clustering: Clustering is the process of partitioning or grouping a given set of patterns into disjoint clusters. In K mean method the number of k clusters, here k is assumed to be fixed. Let the k prototypes \((w_1, \ldots, w_k)\) be initialized to one of the n input patterns \((i_1, \ldots, i_n)\). The steps of clustering technique are:

Each cluster \(C_j\) is associated with prototype \(w_j\)

Repeat
for each input vector \(i_l\) where \(l \in \{1, \ldots, n\}\), do

Assign \(i_l\) to the cluster \(C_j^*\), with nearest prototype \(w_j\) for each cluster \(C_j\), where \(j \in \{1, \ldots, k\}\),

Up date the prototype \(w_j\) to be the Centroid of all samples currently in \(C_j\) so that:

\[
 w_j = \frac{\sum_{i \in C_j} i}{|C_j|} 
\]  

(1)

Compute the error function:

\[
 E = \sum_{i \in C_j} \left| i - w_j \right|^2
\]  

(2)

Until E does not change significantly or cluster membership no longer changes. It shows a high level description of the k-means clustering algorithm for automatic instance clustering and performs correlation analysis between those clusters.

Kernel Canonical Correlation Analysis (KCCA): The goal of Canonical correlation analysis is to find the correlation analysis between two clusters to identify the exact correspondence between those two clusters.

The \(w_x\) for \(x\) and \(w_y\) for \(y\) such that the linear transformation \(x \rightarrow \langle w_x, x \rangle\) and \(y \rightarrow \langle w_y, x \rangle\) applied to the dataset \(D\) results in the maximum correlation. The correlation coefficient is defined as:

\[
 \rho = \max_{w_x, w_y} \frac{w_x^T C_{xx} w_y w_y^T C_{yy} w_x}{\sqrt{w_x^T C_{xx} w_x w_y^T C_{yy} w_y}}
\]  

(3)

subject to:

\[
 w_x^T C_{xx} w_x = 1
\]

\[
 w_y^T C_{yy} w_y = 1
\]

The problem is equivalent to solving a generalized eigenvalue problem, where, \(A = C_{xx}^{-1} C_{xy} C_{yx}^{-1} C_{yy}\) and the solution to \(w_x\) are the eigen-vectors of A. The weights \(w_x\) and \(w_y\) can be expressed in the dual form as follows:

\[
 w_x = \sum_{i=1}^l \alpha_i x_i = \alpha^T D_x^T
\]  

(4)

\[
 w_y = \sum_{i=1}^l \beta_i y_i = \beta^T D_y^T
\]  

(5)
The product of the data matrices can be replaced by kernel matrices $K_x$ and $K_y$:

$$\rho_{\alpha\beta} = \max_{\alpha,\beta} \frac{\alpha^T K_{\alpha} + \alpha^T K_{\alpha} \beta^T K_{\beta}}{\sqrt{\alpha^T K_{\alpha} \alpha + \beta^T K_{\beta} \beta}}$$

$$\alpha^T K_{\alpha} + \alpha^T K_{\alpha} \beta^T K_{\beta} = 1$$

$$\beta^T K_{\beta} + \beta^T K_{\beta} \alpha^T K_{\alpha} = 1$$

After setting up the Lagrangian, it can be shown that the solution to the correlation analysis for $\alpha$ can be obtained by solving the general eigen-value problem. Multiple Instance Learning scenarios assumes that the label $y_i$ is assigned to a bag of instance $X_i = \{x_{i1}, ..., x_{ik}\}$ where $k_i$ is the number of instances in the bag and each instance $x_{ij} \in R^d$ where $d$ is the dimensionality of the feature space used to represent each instance. The bag is assigned a positive label $y_i = 1$ if at least one of the instances in $X_i$ is positive. The bag is assigned a negative label $y_i = -1$ if it only contains only negative instances. Patterns are grouped into bags and a label is attached to each bag and not to every pattern. A set of input patterns $x_1...x_n$ grouped into bags $B_1,..., B_m$ with $B_i = \{x_i; i \in I\}$ for a given index sets $I \subseteq \{1, ..., n\}$.

Initialize $y_i = Y_i$, for $i \in I$

**REPEAT**

- compute SVM solution $w$, $b$ for data with input labels
- compute outputs $f_i = (w, x_i) + b$ for all $x_i$ in positive bags
- set $y_i = \text{sgn}(f_i)$ for every $i \in I$, $Y_i = 1$
- For (every positive bag $B_i$)
  - IF $(\Sigma_{i \in I}(1+y_i))/2 = 0$
  - Compute $I^* = \text{arg max}_{i \in I} f_i$
- Set $y_i^* = 1$
**END**

**WHILE** (input labels have changed)

**OUTPUT** ($w$, $b$)

In multi instance learning the training set comprises the labeled bags that are composed of unlabelled instance and the task is to predict the labels of unseen bags.

**Image classification**: It is used to identify the object classes and image concepts from the test images. It automatically clusters the image instance and performs correlation analysis on different cluster to know exact correspondence between image instances. Multi instance learning is used to improve the classification performance. The test image is first segmented into multiple image regions or multisresolution image patches. Each image region or image patch (image instance) is first classified into the most relevant object class $C_k$ on the visual concept network, which has the maximum value of the confidence score. The confidence scores for the inter-related object classes are further calculated to determine the potential classification paths. If the confidence score for some of these inter-related object classes are above a given threshold $\delta = 0.65$. The classification paths are selected. Otherwise, the relevant classification paths are terminated. The test image is matched with the database and produces the domain specific images. Once a test image is classified, multiple text terms for interpreting the inter-related object classes and image concepts on the visual concept network become multiple labels for annotating its semantics more sufficiently:

For ($x \in R^d$) for $j \in 1,...,|L|$

- do > single-label transformation and training:
  - $D' = \{\}$
  - for $(x, S) \in D$
  - do $D' = D' \cup ((x, l_1, ..., l_j), l_j)$
  - train $C_j$ to predict binary relevance of $l_j$

Training Phase for Dataset $D$ and Label set $L$

Each classifier $C_j$ in the chain is responsible for learning and predicting the binary association of label $l_j$ given the feature space, augmented by all prior binary relevance predictions in the chain $l_1, ..., l_j$.

**EXPERIMENTS AND RESULTS**

The corel dataset is a popular benchmark for image classification. It is based on 5000 Corel images, 4500 of which are used for training and the rest 500 for testing. The Proposed system uses the contourlet transform is used for the initial edge based it improve the performance of classification. It is compared with edge based technique like canny, Sobel, Prewitt, Roberts. The results shown for all edge based approaches are given in Fig. 3. The results

![Fig. 3: (a) Canny (b) Roberts](attachment:image.png)
are shown. In Fig. 4 the performance of SVM classifier is compared based on domain specific images based on precision and recall. It significantly improves the performance of classification accuracy.

The precision and recall values based on domain specific images results are shown in Table 1.

### CONCLUSION

In this study, a Multi instance learning with kernel canonical correlation analysis algorithm is developed to improve the performance of image classification. KCCA algorithm used for identifying correspondence between the image instances. Multi Instance Learning is used for characterizing the inter-concept visual similarity contexts more accurately in the high dimensional multimodal feature space and determining the inter-related learning task directly in feature space.

### REFERENCES


