Transferring CNNs to Multi-instance Multi-label Classification on Small Datasets

Mingzhi Dong[1], Kunkun Pang[2], **Yang Wu**[3], Jinghao Xue[1], Timothy Hospedales[2], Tsukasa Ogasawara[3]

[1]University College London [2] University of Edinburgh
[3]**Nara Institute of Science and Technology**

ICIP, 2017

Outline

1. Multi-instance Multi-Label (MIML) Problems

2. CNNs and MIML Problems

3. Transfer CNNs to small MIML tasks

4. Experiments

Multi-instance Multi-Label (MIML) Problems

- **Multi-label problems**
  Each instance is associated with *a set of labels*.

- **Multi-instance problems**
  The learner receives *a set of bags, each containing many instances*. Instead of receiving the labels of the instances, in multi-instance problems, the bags are labelled.

- **Multi-instance Multi-label (MIML) problems**
  The learner receives *a set of bags* which are associated with *a set of labels*.

- **Image Tagging as a MIML Problem**
  Multi-label: each image is likely worth many words;
  Multi-instance: each image can be viewed as a bag of local regions, the labels are assigned to a whole image but not its specific regions.

Image Tagging as a MIML Problems

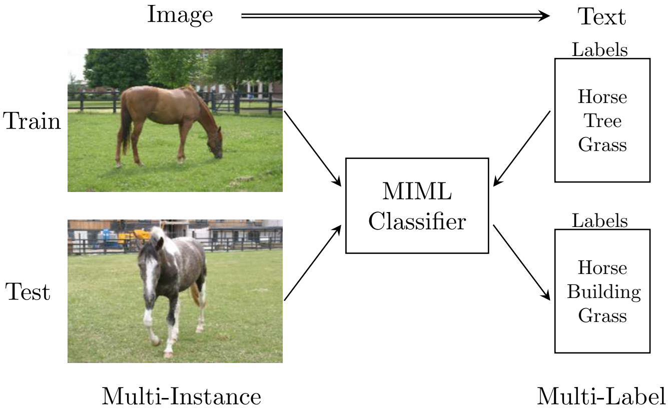

Image → Text

Train

Test

MIML Classifier

Labels

Horse
Tree
Grass

Labels

Horse
Building
Grass

Multi-Instance

Multi-Label

Figure: Multi-instance multi-label learning for image tagging.

Outline

1. Multi-instance Multi-Label (MIML) Problems
2. CNNs and MIML Problems
3. Transfer CNNs to small MIML tasks
4. Experiments

CNNs and MIML Problems

CNNs are suitable for solving MIML problems

- The way to select candidate instances inside a bag (**Mulit-instance**)
  Traditional algorithms: extracting a bag of instances from each image and then applying multi-instance classifiers.
  CNNs: convolutional layers slide through an image (the bag) and create the candidate instances, and the max-pooling layers select the most representative instances inside the bag.

- The way to depict the relationship between the labels (**Multi-label**)
  Traditional algorithms: solve multi-label problems by explicitly or implicitly building the relationship between the labels;
  CNNs: by utilizing a deep hierarchical structure, different levels of representation of all the labels have already been embedded in the network, and the relationship of the labels can be explored via checking the sharing of the neurons.

CNNs and MIML Problems

How to apply CNNs to MIML problems?

- **A simple way** to adapt CNNs for MIML problems
  Change the original multi-class classifier in the Softmax Layer into a multi-label classifier, such as *the multiple binary-class logistic regression (LR)*.

- **Shortage**
  Most existing MIML datasets only have
  a relatively small amount of training data, as the annotation costs for them are generally much higher. Hence, it will be hard to directly train an effective CNN model for such cases.

Outline

1. Multi-instance Multi-Label (MIML) Problems
2. CNNs and MIML Problems
3. Transfer CNNs to small MIML tasks
4. Experiments

Transfer CNNs to small MIML tasks

1. **Extract Features from pre-trained CNN models**
   Based on the VGG 16 layers network (VGG16) pre-trained on Imagenet, we extract features from each group of its layers, which enable a depiction of the *multi-level relationship between the labels*.

2. **L_1-norm regularized Logistic Regression (L_1LR)**
   Then an L_1-norm regularized Logistic Regression (L_1LR) is adopted to learn one classifier for each label. The aim of utilizing the sparsity regularization is to encourage the classifiers to select only much smaller subsets of "effective" features for specific labels.
   Learning the j-th classifier can be expressed as

$$\min_{\boldsymbol{w}_j} \frac{1}{m} \sum_{i=1}^{m} \log(1 + \exp(-y_{ij}(\boldsymbol{x}_i^T \boldsymbol{w}_j + b))) + \lambda |\boldsymbol{w}_j|_1. \qquad (1)$$

Framework Illustration

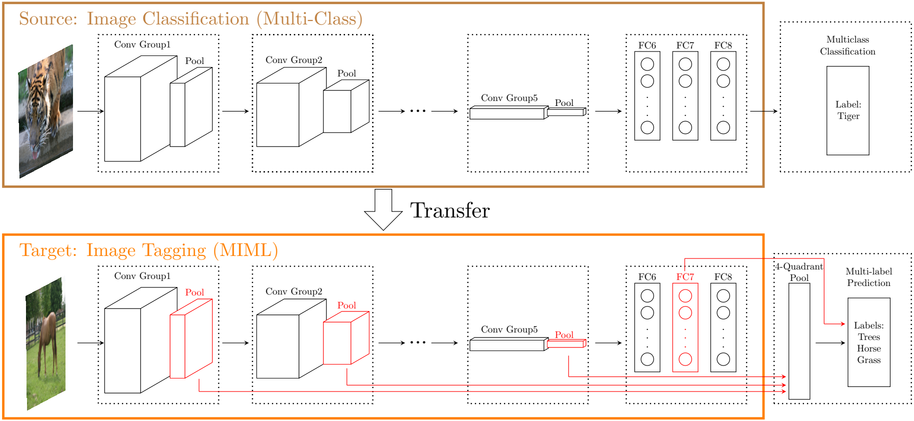

Figure: The framework of Transfer CNN to small MIML tasks. Source Task: Multi-class Classification (Image Classification), outputting one class label; Target Task: MIML Recognition (Image Tagging), outputting multiple labels.

Outline

1. Multi-instance Multi-Label (MIML) Problems

2. CNNs and MIML Problems

3. Transfer CNNs to small MIML tasks

4. Experiments

Settings

- **MIML Datasets**
  MSRC v2, a subset of the Microsoft Research Cambridge (MSRC) image dataset with 591 images and 23 labels (on average 2.5 labels per image),
  Scene, a natural scene dataset with 2000 images and 5 different labels (on average 1.2 labels per image);

- The VGG features are extracted via **Caffe**;

- LR and L_1LR are implemented via **Liblinear**.

- The shrinkage parameter λ of L_1LR is fixed as default of 1.

Results

Table: Performance comparison (mean ± std). The symbol ↑ (↓) indicates that the larger (smaller) the value, the better the performance.

	VGG+L_1LR	**VGG+LR**	MIMLfast	DBA	KISAR
MSRC v2					
a.p. ↑	.933 ± .010	.851 ± .016	.688 ± .017	.326 ± .016	.666 ± .018
co. ↓	.102 ± .006	.140 ± .012	.238 ± .014	.837 ± .018	.254 ± .015
h.l. ↓	.033 ± .003	.051 ± .003	.100 ± .007	.140 ± .006	.086 ± .004
o.e. ↓	.060 ± .017	.149 ± .024	.295 ± .025	.415 ± .026	.341 ± .031
r.l. ↓	.018 ± .003	.045 ± .007	.108 ± .009	.675 ± .017	.131 ± .010
Scene					
a.p. ↑	.948 ± .006	.926 ± .004	.770 ± .015	.600 ± .013	.772 ± .012
co. ↓	.082 ± .006	.096 ± .004	.207 ± .012	.334 ± .011	.204 ± .008
h.l. ↓	.070 ± .004	.090 ± .004	.188 ± .009	.269 ± .009	.194 ± .005
o.e. ↓	.082 ± .010	.114 ± .007	.351 ± .023	.386 ± .025	.351 ± .020
r.l. ↓	.038 ± .005	.055 ± .004	.189 ± .014	.348 ± .012	.185 ± .010

Results

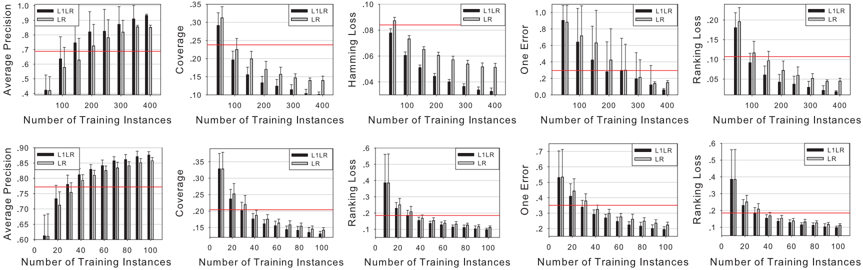

Figure: The performance with a relatively small number of training instances. The red line indicates the best performance of the 6 state-of-the-art algorithms with 2/3 training instances (394 training instances on MSRC dataset on the upper 5 panels, and 1333 training instances on Scene dataset on the lower 5 panels). The performance is based on 30 random repetition and the bar indicates the standard deviation.

Conclusion

- We propose a CNN-based transfer learning framework with sparsity regularization for multi-instance multi-label problems with small datasets;
- The proposal has achieved substantial improvement in classification performance when compared with the state-of-the-art MIML algorithms.